DOCUMENT RESUME

ED 168 543 IR 007 069

AUTHOR	Black, Doris E., Comp.
TITLE	Mathematical and Statistical Software Index. Final Report.
INSTITUTION	Air Force Human Resources Lab., Brooks AFB, Texas.
REPORT NO	AFHRL-TR-78-66
PUB DATE	Nov 78
NOTE	95p.
AVAILABLE FROM	Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20204 (671-056/89)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Analysis of Variance; *Computer Programs; Correlation; Factor Analysis; Item Analysis; *Library Collections; Linear Programing; Multiple Regression Analysis; Statistical Analysis
IDENTIFIERS	*Computer Software; *Scientific and Technical Information

ABSTRACT

Brief descriptions are provided of general-purpose mathematical and statistical software, including 27 "stand-alone" programs, three subroutine systems, and two nationally recognized statistical packages, which are available in the Air Force Human Resources Laboratory (AFHRL) software library. This index was created to enable researchers to identify mathematical or statistical software currently operational on the AFHRL computer system. Information on 27 "stand-alone" programs is presented in the form of an abstract for each item and a list of program documentation references for the reader who wishes to obtain formal program documentation. Each abstract includes general program description, program origin, important program restrictions, and warning about excessive computer execution time incurred in certain circumstances, followed by the symbol of the branch of AFHRL responsible for program development and documentation references. The general descriptions of the software library's three mathematical and/or statistical subroutine systems include a list of the subroutines contained in each system and documentation references. Information provided on two nationally recognized statistical packages, Biomedical Computer Programs and STATJOB, both available in the software library, includes general descriptions, a list and brief descriptions of programs in each package, and package documentation references. (Author/CWM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

AFHRL-TR-78-66

AIR FORCE HUMAN RESOURCES LABORATORY

MATHEMATICAL AND STATISTICAL SOFTWARE INDEX

Compiled by

Doris E. Black

COMPUTATIONAL SCIENCES DIVISION
Brooks Air Force Base, Texas 78235

November 1978

Approved for public release; distribution unlimited.

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This final report was submitted by Computational Sciences Division, under project 6323, with HQ Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235. Doris E. Black (SMS) was the Principal Investigator for the Laboratory.

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

ROBERT A. BOTTENBERG, Chief
Computational Sciences Division

RONALD W. TERRY, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFHRL-TR-78-66
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER

4. TITLE *(and Subtitle)*: MATHEMATICAL AND STATISTICAL SOFTWARE INDEX
5. TYPE OF REPORT & PERIOD COVERED: Final
6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): Doris E. black
8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Computational Sciences Division
Air Force Human Resources Laboratory
Brooks Air Force Base, Texas 78235

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
62703F
63230512

11. CONTROLLING OFFICE NAME AND ADDRESS
HQ Air Force Human Resources Laboratory (AFSC)
Brooks Air Force Base, Texas 78235

12. REPORT DATE: November 1978
13. NUMBER OF PAGES: 94

14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*

15. SECURITY CLASS. *(of this report)*: Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for publice release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

automatic interaction detector
analysis of variance
Biomedical Computer Programs
correlation
cross-classification
curve-fitting
factor analysis
Functional Mathematical Programming System
item analysis
linear programming

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

This report is an abridged documentation source for the Air Force Human Resources Laboratory (AFHRL) mathematical and statistical software library for use by Air Force personnel researchers. It provides a single reference which researchers may quickly scan to identify mathematical or statistical computer software which is currently operational on the AFHRL Univac 1108 computer system and which is available to them for use in their research projects. The report is comprised of four chapters with the first chapter devoted to introductory information; the second chapter, to descriptions of 27 computer programs; the third chapter, to descriptions of the library's subroutine systems; and the final chapter, to the nationally recognized statistical packages available in the software library.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

Item 19 Continued:

mathematical computer software
multiple discriminant analysis
multivariate analysis of variance
multivariate normal probability
PERSUB
regression
regression equation grouping
statistical computer software
STATJOB
test for normality
Univac MATH-PACK
Univac STAT-PACK

PREFACE

The mathematical and statistical software library which is in residence in the Air Force Human Resources Laboratory (AFHRL) Univac 1108 computer system is maintained and undergoes continued development in support of Air Force personnel and training research with the objectives to facilitate and to improve the analysis of human resources data. The compilation of this software index is in line with these objectives making available an abridged documentation source for the software library which will expedite the work of Air Force personnel researchers. Work on the software index was accomplished under project 6323, Personnel Data Analysis; task 632305, Development of Analytic Methodology for Air Force Personnel Research Data.

Many people over a number of years have contributed to the development and implementation of programs in the AFHRL mathematical and statistical software library. While those responsible for the specific programs are identified in their formal program documentation reports, it is appropriate to acknowledge here the key personnel currently responsible for the ongoing development of the software library. They include Dr. Janos B. Koplyay, Mr. C. Deene Gott, Mr. Charles R. Rogers, and Mr. Jacob E. Myer, Jr. Special acknowledgement also goes to Mr. Larry K. Whitehead for his programming efforts in program development, implementation, and maintenance.

TABLE OF CONTENTS

		Page
I.	Introduction	1
II.	Mathematical and Statistical Programs	3
	AID-4 (Automatic Interaction Detector)	3
	ANOVA-UNEQ (Unequal N Analysis of Variance)	5
	BAYS (Bayesian Method for Categorical Prediction)	6
	CANCOR (Canonical Correlation)	8
	COMPARE (Compare One-way Frequencies)	9
	CROSS-CLAS (Cross-Classification)	10
	CURVE-FIT (Fourier-Pearson Curve-Fitting)	11
	DIG (Frequency Distribution Generator)	12
	DIM12 (Multi-variable Frequency Distribution)	13
	DIP (Distribution Print Program)	14
	FMPS (Functional Mathematical Programming System)	15
	GEN-HIER-GRP (Generate HIER-GRP Input Data)	16
	GUTTMAN (Guttman Item Analysis)	17
	HIER-GRP (Regression Equation Grouping)	18
	HITAB$ (Hit-tables for Dichotomous Criteria)	20
	IAP (Item Analysis of Achievement Type Tests)	21
	IAPG (Analysis of Questionnaire Type Tests)	22
	IAPG-HITS (Hit-table for Dichotomous Criteria)	23
	MANOVA (Multivariate Analysis of Variance)	24
	MAX-FACTOR (Factor Analysis)	25
	MEANSD (Mean & Standard Deviation Program)	26
	MULT-DISCR (Multiple Discriminant Analysis)	27
	MULT-PROB (Multivariate Normal Probability)	28
	REGREX (Stepwise Regression)	29
	TRICOR (Correlation and Regression Program)	30
	WPRIME (Residuals Test for Normality)	31
	1WAY (One-way Frequency Distribution Print)	32
III.	Mathematical and Statistical Subroutine Systems	33
	Univac MATH-PACK	34
	Univac STAT-PACK	38
	PERSUB System	42
IV.	Nationally Recognized Statistical Packages	46
	Biomedical Computer Programs	47
	STATJOB	74
References		83
Bibliography		87
Technical Notes		88

MATHEMATICAL AND STATISTICAL SOFTWARE INDEX

1. INTRODUCTION

THE COMPUTATIONAL SCIENCES DIVISION (SM) OF THE AIR FORCE HUMAN RESOURCES LABORATORY (AFHRL) IS RESPONSIBLE FOR CONDUCTING RESEARCH AND DEVELOPMENT IN THE AREAS OF STATISTICS, MATHEMATICS, COMPUTER OPERATIONS, AND RESEARCH DESIGN IN SUPPORT OF AIR FORCE PERSONNEL AND TRAINING RESEARCH WITH THE OBJECTIVES OF FACILITATING AND IMPROVING THE ANALYSIS OF HUMAN RESOURCES DATA. AS A RESULT OF THIS MISSION, A SIZEABLE MATHEMATICAL AND STATISTICAL SOFTWARE LIBRARY WHICH ENCOMPASSES A BROAD RANGE OF METHODOLOGY IS RESIDENT IN THE AFHRL UNIVAC 1108 COMPUTER SYSTEM HOUSED IN SM. PAST SM EXPERIENCE WITH AIR FORCE PERSONNEL RESEARCHERS HAS REVEALED A NEED FOR AN ABRIDGED DOCUMENTATION SOURCE ON THE SOFTWARE LIBRARY WHICH RESEARCHERS MAY QUICKLY SCAN TO IDENTIFY THE MATHEMATICAL OR STATISTICAL SOFTWARE WHICH IS CURRENTLY OPERATIONAL ON THE AFHRL COMPUTER SYSTEM AND IS AVAILABLE TO THEM FOR USE IN THEIR RESEARCH PROJECTS. THE PURPOSE OF THIS REPORT IS TO FILL THIS NEED BY PROVIDING A SOFTWARE INDEX WHICH CONTAINS BRIEF DESCRIPTIONS OF GENERAL-PURPOSE MATHEMATICAL AND STATISTICAL SOFTWARE IN THE AFHRL SOFTWARE LIBRARY.

THE SOFTWARE DESCRIPTIONS PROVIDED IN THIS REPORT ARE ORGANIZED INTO THREE CHAPTERS (I.E., CHAPTERS 2 THROUGH 4). CHAPTER 2 DEALS WITH THE "STAND-ALONE" PROGRAMS IN THE SOFTWARE LIBRARY, WHILE CHAPTER 3 DESCRIBES THE LIBRARY'S SUBROUTINE SYSTEMS. THE FINAL CHAPTER, CHAPTER 4, DISCUSSES THE NATIONALLY RECOGNIZED STATISTICAL PACKAGES WHICH ARE ALSO AVAILABLE IN THE SOFTWARE LIBRARY.

CHAPTER 2 CONTAINS INFORMATION ON 27 PROGRAMS, WITH THE INFORMATION FOR EACH PROGRAM BEING COMPRISED OF TWO PARTS: AN ABSTRACT AND A LIST OF PROGRAM DOCUMENTATION REFERENCES FOR THE READER WHO WISHES TO OBTAIN THE FORMAL DOCUMENTATION FOR THE PROGRAM. IN EACH CASE, THE INFORMATION INCLUDED IN THE ABSTRACT FOR A PROGRAM GIVES A GENERAL DESCRIPTION OF THE PROGRAM; THE ORIGIN OF THE PROGRAM, IF THE PROGRAM WAS OBTAINED FROM A SOURCE OUTSIDE OF AFHRL; IMPORTANT PROGRAM RESTRICTIONS (E.G., MAXIMUM NUMBER OF VARIABLES PERMITTED); AND A WARNING (IF APPLICABLE) ABOUT EXCESSIVE COMPUTER EXECUTION TIME WHICH MIGHT BE INCURRED IN CERTAIN CIRCUMSTANCES. FOLLOWING THE ABSTRACT IN A SET OF PARENTHESES IS THE OFFICE SYMBOL OF THE SM BRANCH RESPONSIBLE FOR THE PROGRAM (I.E., "SMA" FOR THE ANALYSIS AND PROGRAMMING BRANCH, "SMO" FOR THE COMPUTER OPERATIONS BRANCH, AND "SMS" FOR THE STATISTICAL AND COMPUTER TECHNOLOGY BRANCH). THE DOCUMENTATION REFERENCE FOR EACH PROGRAM IS USUALLY IN THE FORM OF A TECHNICAL REPORT OR A COMPUTERIZED DOCUMENTATION REPORT PRODUCED BY THE UNIVAC 1108 DOC PROCESSOR. FOR THE READER WITH ACCESS TO THE AFHRL UNIVAC 1108, THE DOCUMENTATION REFERENCE ALSO CONTAINS (IF APPLICABLE) THE RETRIEVAL COMMAND WHICH CAN BE EXECUTED TO SECURE

A COPY OF THE DOC PROCESSOR REPORT. READERS WHO DO NOT HAVE ACCESS TO THE AFWRL COMPUTER AND WISH TO OBTAIN A COPY OF THE DOC PROCESSOR REPORT MAY ADDRESS THEIR REQUESTS TO AFWRL/SN, BROOKS AIR FORCE BASE, TEXAS 78235. (NOTE: NON-AIR FORCE REQUESTERS WILL BE REQUIRED TO SECURE RELEASE APPROVAL FOR THE REQUESTED SOFTWARE DOCUMENTATION FROM THE PROGRAMS AND TECHNOLOGY DIVISION OF DATA AUTOMATION, HEADQUARTERS AIR FORCE SYSTEMS COMMAND (HQ AFSC/ADCT).)

CHAPTER 3 PROVIDES INFORMATION ON THE SOFTWARE LIBRARY'S THREE MATHEMATICAL AND/OR STATISTICAL SUBROUTINE SYSTEMS. A GENERAL DESCRIPTION OF EACH SUBROUTINE SYSTEM IS GIVEN ALONG WITH A LIST OF THE SUBROUTINES CONTAINED IN EACH SYSTEM AND A LIST OF DOCUMENTATION REFERENCES.

CHAPTER 4 OF THIS REPORT IDENTIFIES THE TWO NATIONALLY RECOGNIZED STATISTICAL PACKAGES WHICH ARE AVAILABLE IN THE SOFTWARE LIBRARY. THE INFORMATION PROVIDED ON EACH PACKAGE INCLUDES A GENERAL DESCRIPTION OF THE PACKAGE, A LIST AND BRIEF DESCRIPTIONS OF THE PROGRAMS CONTAINED IN THE PACKAGE, AND A LIST OF REFERENCES TO THE PACKAGE DOCUMENTATION.

2. Mathematical and Statistical Programs

2.1. AID-4 (Automatic Interaction Detector)

2.1.1. Abstract

AID-4 is a computer program written primarily in Fortran and adapted from an earlier program (Sonquist and Morgan, 1964) obtained from the University of Michigan. It is useful in the automatic identification of interactions among predictor variables for multiple linear regression analysis and is essentially a model identification process since it provides an optimum solution for the analysis. Given an initial group of data, the definition of the criterion variable, and the definition of the predictor variables, the basic idea of the AID-4 algorithm is to explain the variance of the criterion variable by the sequential splitting of the original group into subgroups. The splitting is done in such a way as to minimize the within groups (error) sum of squares. This is accomplished by the examination of each possible split of every predictor variable of the current candidate group to be split and by the selection of the split giving the smallest within groups sum of squares. Three basic statistics are reported at each split: an R-squared value indicating the percentage of the criterion variance explained through the current splitting state; an F-value indicating the significance of the reduction in the error sum of squares due to the current split; and an F-value for a one-way analysis of variance considering all groups at this stage. AID-4 will handle up to 300 predictor variables which may be either categorical or continuous. All predictors are recoded to categorical variables according to user specifications. The maximum number of recode categories for a predictor is 50. (SMS)

2.1.2. Documentation References

Kaplan, D.R., Cott, L.D., & Elton, J.M. Automatic Interaction Detector-Version 4 (AID-4) Reference Manual. AFHRL-TR-73-17, AD-773 803. Lackland AFB, Tx: Personnel Research Division, Air Force Human Resources Laboratory, October 1973.

[illegible]

[illegible]

[illegible]

[redacted]

[redacted]

[redacted]

2.3. BAYS (BAYESIAN METHOD FOR CATEGORICAL PREDICTION)

2.3.1. ABSTRACT

WITH A SET OF CATEGORICAL PREDICTORS AND A CATEGORICAL CRITERION AS INPUT VARIABLES, THE OBJECTIVE OF THE ANALYSIS CARRIED OUT BY THE BAYS PROGRAM IS TO FIND A SUBSET OF THE PREDICTORS WHICH MOST EFFECTIVELY DISCRIMINATES AMONG THE CRITERION CATEGORIES. BAYS IS BASED ON A MODIFIED VERSION OF THE "ABCD TECHNIQUE" DESCRIBED BY MOONAN (1972) AND USES BAYES' FORMULA TO COMPUTE PROBABILITIES OF CLASS MEMBERSHIP FOR EACH CASE WITH THE RESULT THAT AN INDIVIDUAL IS ASSIGNED TO THE CRITERION CATEGORY FOR WHICH HIS A POSTERIORI PROBABILITY IS HIGHEST OR HIS COST OF MISCLASSIFICATION IS LOWEST (AT THE OPTION OF THE USER). AN IMPROVEMENT TO THE ABCD TECHNIQUE IS EMPLOYED BY BAYS IN WHICH A STEPWISE PROCEDURE IN THE MODEL-BUILDING ALGORITHM CAN CAUSE VARIABLES TO BE ELIMINATED AFTER THEY HAVE BEEN ADDED TO THE PREDICTIVE SCHEME. AT EACH STAGE OF THE MODEL-BUILDING PROCESS, THE PREDICTIVE COMPOSITE IS FORMED WHICH CORRESPONDS TO THE HIGHEST CLASSIFICATION ACCURACY RESULTING FROM ALL POSSIBLE ADDITIONS (OR DELETIONS) OF ONE VARIABLE TO (OR FROM) THE PREDICTIVE COMPOSITE EXISTING AT THE PREVIOUS STAGE. BY REVIEWING THE HIT TABLE OUTPUT PRODUCED AT EACH STAGE, A USER CAN COMPARE THE CLASSIFICATION ACCURACIES OF THE PREDICTIVE COMPOSITES AS THEY ARE GENERATED. ADDITIONALLY, PREDICTIVE COMPOSITES CAN BE EVALUATED IN TERMS OF MISCLASSIFICATION COSTS. THE FOLLOWING OPTIONS ARE AVAILABLE TO THE BAYS USER: (A) EMPIRICAL PROBABILITIES MAY BE SUPPLIED BY THE USER OR THE PROGRAM WILL COMPUTE THEM FROM THE APPROPRIATE SAMPLE(S); (B) THE ITEMS MAY BE SELECTED SO AS TO MAXIMIZE CORRECT CLASSIFICATIONS OR MINIMIZE EXPECTED COSTS OF MISCLASSIFICATIONS; (C) CERTAIN VARIABLES MAY BE FORCED INTO THE PREDICTIVE SCHEME INITIALLY; AND (D) THE SET OF PREDICTORS SELECTED MAY BE CROSS-VALIDATED ON UP TO TWO SAMPLES WITHIN THE SAME RUN.

BAYS IS PROGRAMMED TO HANDLE UP TO 200 ITEMS, FIVE CRITERION CATEGORIES, AND THREE SAMPLES, WITH THE LIMITATION THAT EXCESSIVE COMPUTER RUN TIMES CAN BE INCURRED FOR PROBLEMS INVOLVING LARGE SAMPLES. (SMS)

MATHEMATICAL AND STATISTICAL SOFTWARE INDEX

2.3.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - BAYS.

BAYS REPORT RETRIEVAL COMMAND: @T*T.DOC BAYS.

2.9. CANCOR (CANONICAL CORRELATION)

2.9.1. ABSTRACT

THE PURPOSE OF CANONICAL ANALYSES IS TO DEFINE THE PRIMARY INDEPENDENT DIMENSIONS (BASIS VECTORS) WHICH RELATE ONE SET OF VARIABLES TO ANOTHER SET OF VARIABLES. THE TECHNIQUE IS PRIMARILY DESCRIPTIVE, ALTHOUGH THE METHOD INVOLVES FINDING SETS OF WEIGHTS WHICH MAXIMIZE THE CORRELATIONS BETWEEN TWO COMPOSITE VARIABLES (ONE FOR EACH SET OF ORIGINAL VARIABLES). THE TWO OUTPUTS SHOULD SUGGEST ANSWERS TO QUESTIONS CONCERNING THE NUMBER OF WAYS THE TWO SETS OF MEASURES ARE RELATED, THE STRENGTHS OF THE RELATIONSHIPS, AND THE NATURE OF THE RELATIONSHIPS SO DEFINED.

THIS CANONICAL CORRELATION ROUTINE IS A MODIFICATION AND COMBINATION OF THE ROUTINES IN "MULTIVARIATE PROCEDURES FOR THE BEHAVIORAL SCIENCES" BY COOLEY AND LOHNES (1962) AND "FORTRAN PROGRAMMING FOR THE BEHAVIORAL SCIENCES" BY VELDMAN (1967). THE PROGRAM COMPUTES A FULL SET OF CANONICAL CORRELATIONS FOR TWO SETS OF VARIABLES. THE OUTPUT CONTAINS MEANS AND STANDARD DEVIATIONS FOR ALL VARIABLES; SET A CORRELATIONS; SET B CORRELATIONS; CROSS CORRELATIONS (BETWEEN THE A AND B VARIABLES); WILKS' LAMBDA AND CHI-SQUARE FOR THE FULL SET OF ROOTS. ADDITIONALLY, FOR EACH INDIVIDUAL ROOT, THE PROGRAM REPORTS THE CHI-SQUARE VALUE, THE DEGREES OF FREEDOM, THE SIGNIFICANCE OF THE CHI-SQUARE, BOTH SETS OF WEIGHTS, AND (OPTIONAL) CORRELATIONS BETWEEN THE ORIGINAL A AND B VARIABLES AND THE CANONICAL VARIATES.

THE PROGRAM IS RESTRICTED TO A MAXIMUM OF 50 VARIABLES PER SET. (SMS)

2.9.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - CANCOR.

CANCOR REPORT RETRIEVAL COMMAND: @T.T.DOC CANCOR.

2.5. COMPARE (COMPARE ONE-WAY FREQUENCIES)

2.5.1. ABSTRACT

THIS PROGRAM COMPARES ONE-WAY FREQUENCY DISTRIBUTIONS FROM DIFFERENT INPUT FILES. THE DISTRIBUTIONS ARE GENERATED BY THE DIG PROGRAM. (SEE SECTION 2.8 FOR A DESCRIPTION OF THE DIG PROGRAM.) COMPARE IS ESPECIALLY APPROPRIATE FOR COMPARISON OF RESPONSES TO SIMILAR QUESTIONS FROM DIFFERENT SURVEY DATA SETS.

UP TO 9 FREQUENCY DISTRIBUTION FILES MAY BE INPUT TO A SINGLE COMPARE RUN. FREE FORMAT CONTROL COMMANDS DIRECT THE WAY IN WHICH THE PROGRAM WILL ORGANIZE THE OUTPUT REPORT.

ALL PRINTED REPORTS PRODUCED BY COMPARE CONFORM TO THE AFHRL/SM STANDARDS FOR FINAL REPORT PRODUCTS. (SMA)

2.5.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - SMAZ UTILITY PROGRAMS & PROCESSORS.

REPORT RETRIEVAL COMMAND: @Z.Z.DOC.

2.6. CROSS-CLAS (CROSS-CLASSIFICATION)

2.6.1. ABSTRACT

THE CROSS-CLAS PROGRAM PRODUCES CROSS-CLASSIFICATIONS (ALSO CALLED CROSS-TABULATIONS OR CONTINGENCY TABLES) IN TWO, THREE, OR FOUR DIMENSIONS. ROW AND COLUMN PERCENTAGES, CHI-SQUARES, AND KENDALL'S TAU MAY BE OBTAINED FOR ALL TABLES IF DESIRED.

A MAXIMUM OF 99 PROBLEMS MAY BE EXECUTED IN ONE RUN, EACH PROBLEM CAN PRODUCE UP TO 80 TABLES FROM A GIVEN SET OF CONTROL CARDS AND DATA. A GIVEN SET OF DATA MAY CONSIST OF A MAXIMUM OF 99,999 OBSERVATIONS, WITH NO MORE THAN 80 VARIABLES PER OBSERVATION. DATA MAY BE PROCESSED AS READ IN, OR THEY MAY, IF SPECIFIED, BE RECODED ACCORDING TO VARIOUS SCHEMATA. (SMS)

2.6.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - CROSS-CLAS.

CROSS-CLAS REPORT RETRIEVAL COMMAND: @T.T.DOC CROSS-CLAS.

2.7. CURVE-FIT (FOURIER-PEARSON CURVE-FITTING)

2.7.1. ABSTRACT

CURVE-FIT CONSISTS OF TWO ROUTINES FOR FITTING A CONTINUOUS DISTRIBUTION TO EMPIRICAL STATISTICAL DATA FOR WHICH THE PARENT POPULATION IS UNKNOWN. THESE ROUTINES WERE DEVELOPED BY JAMES A. CHISMAN AT CLEMSON UNIVERSITY IN 1968. ONE ROUTINE FITS A FINITE FOURIER SERIES TO THE SAMPLE POINTS. THE OTHER ROUTINE FITS EACH OF THE 12 PEARSON DISTRIBUTION TYPES (PLUS THE NORMAL DISTRIBUTION) TO A SET OF DATA.

THE FOURIER ROUTINE ACCOMPLISHES FOURIER CURVE-FITTING FOR UP TO 50 EQUALLY SPACED POINTS AND WILL PROVIDE THE USER WITH THE FOLLOWING: THE CHI-SQUARE TEST FOR GOODNESS OF FIT; VARIOUS TABLES, INCLUDING A TABLE OF CUMULATIVE PROBABILITIES; AN OPTIONAL GRAPH OF THE DATA POINTS AND THE FITTED CURVE; AND APPROPRIATE DIAGNOSTIC MESSAGES.

THE PEARSON ROUTINE COMPUTES THE PARAMETERS OF EACH OF 13 PEARSON STATISTICAL DISTRIBUTION TYPES FROM THE FIRST FOUR MOMENTS OF THE DATA AS DESCRIBED IN "FREQUENCY CURVES AND CORRELATION" BY W.P. ELDERTON (1953). A CHI-SQUARE TEST IS PROVIDED TO COMPARE THE FIT OF THE VARIOUS DISTRIBUTION TYPES ON A GIVEN SET OF DATA. A MAXIMUM OF 50 INPUT POINTS IS ALLOWED WITH THE REQUIREMENT THAT THEY BE EQUALLY SPACED. (SMS)

2.7.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - CURVE-FIT.

CURVE-FIT REPORT RETRIEVAL COMMAND: @T.T.DOC CURVE-FIT.

2.8. DIG (FREQUENCY DISTRIBUTION GENERATOR)

2.8.1. ABSTRACT

DIG PROVIDES A FAST AND EFFICIENT METHOD OF DISTRIBUTING INFORMATION ON A FILE. DIG IS ACTUALLY A PROGRAM GENERATOR WHICH WILL ACCEPT MINIMAL CONTROL INFORMATION AND IN TURN AUTOMATICALLY GENERATE INTERNAL CODE AND EXECUTE THE CREATED PROGRAM. NOTABLE POINTS:

A. MAY HANDLE 500-600 COMPOSITE FIELDS IN A SINGLE PASS. MAXIMUM FIELD LENGTH IS 30 FIELDATA OR 20 ASCII CHARACTERS.

B. THE BASIC ALGORITHM IS COMPLEX, USING NESTED BINARY TREES FOR INTERNAL STORAGE AND AN ASYNCHRONOUS MERGE (ORDER 16) ON EXTERNAL STORAGE.

C. THE PROGRAM IS EXTREMELY RELIABLE AND ITS EFFICIENCY USUALLY WILL EXCEED THAT OF A SPECIALIZED PROGRAM TO DO THE SAME JOB.

D. 95% OF THE PROGRAMMING WORK IS DONE BY THE PROGRAM.

E. THE DIG FREQUENCY OUTPUT FILE MAY BE REPORTED USING SEVERAL PROGRAMS (DIP, DIM12, JWAY, ETC.).

F. CAN HANDLE AN INPUT BLOCKSIZE UP TO 2049 WORDS.

G. 21 DISTRIBUTIONS OF 1,072,328 CASES (6 REELS) USING 55 FIELDS OF AN AVERAGE LENGTH OF 3 FIELDATA CHARACTERS TOOK 3 HOURS 19 MINUTES OF WALL CLOCK TIME WITH APPROXIMATELY 1 HOUR 18 MINUTES OF CPU TIME.

(SMO)

2.8.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - DISTRIBUTION GENERATOR, DIG.

DIG REPORT RETRIEVAL COMMAND: @S.S.DOC DIG.

2.9. DIM12 (MULTI-VARIABLE FREQUENCY DISTRIBUTION)

2.9.1. ABSTRACT

DIM12 IS A LARGE MODULAR PROGRAM WHICH ENABLES THE PRINTING OF MULTI-VARIABLE DISTRIBUTIONS, INCLUDING FREQUENCY, PERCENTAGE, MEAN, STANDARD DEVIATION, MEDIAN, CUMULATIVE FREQUENCY, AND CUMULATIVE PERCENTAGE DISTRIBUTIONS AS DESIRED. THE MODULARITY OF THE PROGRAM ALLOWS FOR THE FOLLOWING:

A. USE OF LESS CORE, LEAVING MORE ROOM FOR OTHER RUNS IN THE SYSTEM;

B. THE ABILITY TO EASILY STOP AND RESTART THE PROGRAM WITH MINIMUM RECOVERY EFFORT IN A SYSTEM CRASH; AND

C. EASY ADDITION OF NEW MODULES.

RESULTS ARE PRINTED AS A SERIES OF MATRICES WITH PROVISION FOR UP TO 450 COLUMNS (10 COLUMNS PER PAGE) AND 3000 ROWS (DOUBLE OR SINGLE SPACED). EACH MATRIX CAN BE ASSOCIATED WITH ANY COMBINATION OF UP TO 10 MAJOR VARIABLES, WHERE EACH MAJOR VARIABLE IS FIXED AT ANY OF 300 POSSIBLE VALUES. AN ADDITIONAL FEATURE PERMITS THE USER TO RECODE VALUES ON ANY ROW VARIABLE AND TO FORCE THE COLLAPSE OF DISTINCT VALUES ON THE ROW VARIABLE INTO A SINGLE VALUE. ENGLISH LANGUAGE DESCRIPTIONS OF CODED VARIABLES MAY BE USED TO INCREASE THE READABILITY OF THE TABLES. THE USER MAY SPECIFY COMPUTATION OF SUBTOTAL ROWS AT DESIGNATED POSITIONS IN THE MATRIX. THE DATA DISPLAYED FROM THE CHI-SQUARE TEST OF INDEPENDENCE INCLUDE THE CHI-SQUARE VALUE, SIGNIFICANCE LEVEL, DEGREES OF FREEDOM, AND CONTINGENCY COEFFICIENT. (SMA)

2.9.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - FREQUENCY DISTRIBUTION PRINT PROGRAM.

REPORT RETRIEVAL COMMAND: @Z*Z.DOC.

2.10. DIP (DISTRIBUTION PRINT PROGRAM)

2.10.1. ABSTRACT

DIP IS A PRINT ROUTINE FOR THE OUTPUT OF A FREQUENCY DISTRIBUTION GENERATING PROGRAM (E.G., DIG). DIP PACKS THE MAXIMUM NUMBER OF COLUMNS AND ROWS PER PAGE BASED ON THE FIELD SIZE AND AUTOMATICALLY DISPLAYS SUBTOTALS FOR EACH LEVEL OF INTERMEDIATE AND MAJOR VARIABLES WITHIN MULTI-VARIATE DISTRIBUTION REPORTS. A TEST ON THE RESULTS OF 21 DISTRIBUTIONS OF 1,072,328 CASES AS PRODUCED BY DIG TOOK APPROXIMATELY 13 MINUTES OF WALL CLOCK TIME AND ABOUT 6 MINUTES OF CPU TIME TO PRINT 1033 PAGES. THE PRINT WAS 4 TO 5 COLUMNS PER PAGE (ABOUT 200-250 VALUES PER PAGE). (SMO)

2.10.2. DOCUMENTATION REFERENCES

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - DISTRIBUTION PRINT PROGRAM, DIP.

DIP REPORT RETRIEVAL COMMAND: @S.S.DOC DIP.

2.11. FMPS (FUNCTIONAL MATHEMATICAL PROGRAMMING SYSTEM)

2.11.1. ABSTRACT

FMPS IS A MATHEMATICAL PROGRAMMING SYSTEM SUPPLIED BY SPERRY-UNIVAC FOR USE ON THE UNIVAC 1108. THE PROGRAM PROVIDES ALL THE PROCEDURES COMMONLY USED TO SOLVE LINEAR PROGRAMMING PROBLEMS AND HAS THE CAPABILITY TO SOLVE LARGE PROBLEMS. ADDITIONALLY, THE PROGRAM HAS THE FACILITY TO PRODUCE USER-DESIGNED OPTIMIZATION PROCEDURES THROUGH THE USE OF A USER-ORIENTED CONTROL LANGUAGE WHICH RESEMBLES THE FORTRAN LANGUAGE. (SMS)

2.11.2. DOCUMENTATION REFERENCES

UNIVAC SYSTEMS REFERENCE LIBRARY. SPERRY UNIVAC 1100 SERIES FUNCTIONAL MATHEMATICAL PROGRAMMING SYSTEM (FMPS) PROGRAMMER REFERENCE. UP-8198. SPERRY RAND CORPORATION. 1975.

AFWAL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE = FMPS.

FMPS REPORT RETRIEVAL COMMAND: @T*T.DOC FMPS.

2.12. GEN-MIER-GRP (GENERATE MIER-GRP INPUT DATA)

2.12-1. ABSTRACT

GEN-MIER-GRP IS A UTILITY PROGRAM WHICH WILL GENERATE MIER-GRP (SEE SECTION 2.14 FOR A DESCRIPTION OF MIER-GRP) CONTROL CARDS AND DATA USING THE INFORMATION CONTAINED ON A TRICOR (CORRELATION AND REGRESSION PROGRAM) REGRESSION OUTPUT FILE. (SMS)

2.12-2. DOCUMENTATION REFERENCE

AFWAL UNIVAC 1100 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - GEN-MIER-GRP.

GEN-MIER-GRP REPORT RETRIEVAL COMMAND: @T*T.DOC GEN-MIER-GRP.

2.13. GUTTMAN (GUTTMAN ITEM ANALYSIS)

2.13.1. ABSTRACT

THE GUTTMAN ITEM ANALYSIS PROGRAM IS AN ADAPTATION OF THE BASIC PROGRAM OBTAINED BY DR. MALCOLM PEE (APRIL/PEE) FROM THE RESEARCH OFFICE OF THE SCHOOL DISTRICT OF PHILADELPHIA. GUTTMAN'S SCALING TECHNIQUE IS DESIGNED TO BE USED WITH MULTIPLE-CHOICE TESTS FOR WHICH THERE ARE NO PREDETERMINED CORRECT RESPONSES, AND THEREFORE, NO GUIDELINES FOR DIFFERENTIALLY WEIGHTING THE ITEM RESPONSE ALTERNATIVES. THIS PROGRAM USES A MODIFIED VERSION OF THE GUTTMAN WEIGHTING PROCEDURE KNOWN AS THE METHOD OF RECIPROCAL AVERAGES (RAVE). THIS PROCEDURE, WHICH WAS IMPLEMENTED BY BARCO AND MAGSDALE (1964) EMPLOYS AN ITERATIVE APPROACH IN WHICH PREDETERMINED (INITIAL) WEIGHTS FOR THE ITEM RESPONSE CATEGORIES ARE USED AS A SCALING KEY TO COMPUTE A TOTAL SCORE FOR EACH SUBJECT. THESE SCORES AND THE ITEM RESPONSE CATEGORIES ARE THEN USED TO DERIVE A NEW SET OF ITEM OPTION WEIGHTS, WHICH ARE EMPLOYED TO COMPUTE TOTAL SCORES FOR EACH SUBJECT. THIS PROCEDURE MAY BE REPEATED UNTIL SUCCESSIVE SETS OF ITEM OPTION WEIGHTS DO NOT DIFFER APPRECIABLY. THE ULTIMATE SET OF WEIGHTS MAXIMIZES THE INTERNAL CONSISTENCY INDEX OF TEST RELIABILITY FOR THE GROUP OF SUBJECTS ON THAT INSTRUMENT. ITERATIVE OUTPUT INCLUDES THE NEW ITEM OPTION WEIGHTS, MEAN SQUARES FOR ERROR AND SUBJECT VARIANCE, HOYT'S INTERNAL CONSISTENCY COEFFICIENT, OPTIONAL TABLES OF THE DIFFERENCE BETWEEN SETS OF WEIGHTS ON SUCCESSIVE ITERATIONS, AND OPTIONAL VALIDITY COEFFICIENTS BETWEEN THE EXTERNAL (AND/OR INTERNAL) CRITERION AND THE ITERATION SCORES. (SAS)

2.13.2. DOCUMENTATION REFERENCE

APRIL UNIVAC 1108 SOC PROCESSOR REPORT (UNPUBLISHED): TITLE = GUTTMAN.

GUTTMAN REPORT RETRIEVAL COMMAND: @T=?.DOC GUTTMAN.

제2절 [illegible]

제2관 [illegible]

[illegible]

제3관 [illegible]

[illegible]

제4관 [illegible]

[illegible]

[illegible]

[illegible]

[illegible]

2.15. MITABS (HIT-TABLES FOR DICHOTOMOUS CRITERIA)

2.15.1. ABSTRACT

MITABS IS A UTILITY PROGRAM USED TO SUPPLEMENT THE TRICOR CORRELATION AND REGRESSION PROGRAM BY DETERMINING THE ACCURACY OF PREDICTION OF DICHOTOMOUS CRITERIA. THIS IS ACCOMPLISHED BY COMPUTING BINARY CLASSIFICATION TABLES FOR THE DICHOTOMOUS (1/0) CRITERIA USING SPECIAL VERSIONS OF THE PERSUB TABDEV AND HITS SUBROUTINES (SEE SECTION 3.3 FOR A DESCRIPTION OF PERSUB, TABDEV, AND HITS). THE INPUT FILES ARE PREDICTED SCORE OUTPUT FILES AS GENERATED BY THE TRICOR PROGRAM FOR EITHER THE VALIDATION (I.E. REGRESSION) SAMPLE OR BOTH THE VALIDATION AND THE CROSS-VALIDATION SAMPLES. IF THE CROSS-VALIDATION PREDICTED SCORES FILE IS ALSO USED, THE PROGRAM WILL APPLY THE CUTOFF SCORE ESTABLISHED BY THE VALIDATION SCORES TABLE DEVELOPMENT TO THE CROSS-VALIDATION SCORES. (S=S)

2.15.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - MITABS.

MITABS REPORT RETRIEVAL COMMAND: @T*T.DOC MITABS.

2.16. IAP (ITEM ANALYSIS OF ACHIEVEMENT TYPE TESTS)

2.16.1. ABSTRACT

IAP IS AN ITEM ANALYSIS COMPUTER PROGRAM FOR MULTIPLE-CHOICE ACHIEVEMENT TYPE TESTS WITH UP TO 200 ITEMS. THE BASIC CONCEPT IN IAP MAKES USE OF THE ITEM CHARACTERISTIC CURVE AND ITS ASSOCIATED PARAMETERS, XSO AND BETA (TUCKER, 1946; BAKER, 1965). THE ANALYSES PERFORMED BY IAP ENABLE THE USER TO (1) CHOOSE ITEMS WHICH HAVE OPTIMUM DISCRIMINATION POWER AT A CERTAIN ABILITY LEVEL; (2) SCREEN OFF A CERTAIN PERCENTAGE OF A GROUP OF EXAMINEES (3) ESTIMATE THE TRUE SCORE OF AN INDIVIDUAL; AND (4) COMPUTE THE PROBABILITY OF A CORRECT RESPONSE. THE CRITERION UPON WHICH THE PROGRAM BASES ALL THE STATISTICAL ANALYSES MAY BE EITHER THE TOTAL TEST SCORE, WHICH WILL BE CORRECTED FOR GUESSING IF DESIRED, OR A USER-SPECIFIED CRITERION. THE TEST MAY BE TREATED EITHER AS A POWER TEST, IN WHICH CASE THE ANALYSIS OF EACH ITEM IS BASED ON THE TOTAL SAMPLE, OR AS A SPEED TEST, IN WHICH CASE ONLY THOSE REACHING A PARTICULAR ITEM WILL BE CONSIDERED IN THE ANALYSIS OF THAT ITEM. IN ADDITION, THE USER MAY REQUEST A FACTOR ANALYSIS OF THE TETRACHORIC INTER-ITEM CORRELATION MATRIX. THE PRINTED OUTPUT INCLUDES FOR EACH ITEM ALTERNATIVE THE PROPORTION OF SUBJECTS CHOOSING THE ALTERNATIVES; MEANS; STANDARD DEVIATIONS; BISERIAL AND POINT BISERIAL CORRELATIONS WITH STATISTICAL TESTS OF SIGNIFICANCE; ITEM DIFFICULTY; DISCRIMINATION INDEX (BETA); ABILITY LEVEL AT WHICH THE ITEM DISCRIMINATES (XSO); FREQUENCY DISTRIBUTION; PHI COEFFICIENTS (OPTIONAL); IDENTIFICATION OF TOO EASY OR TOO DIFFICULT ITEMS (AS SPECIFIED BY THE USER); AND PLOTS OF THE ITEM CHARACTERISTIC CURVES (OPTIONAL). OPTIONS ARE ALSO AVAILABLE TO CORRECT FOR ITEM/TEST CORRELATION OVERLAP AND FOR SCORING ITEMS HAVING MORE THAN ONE ALTERNATIVE DESIGNATED AS THE CORRECT ANSWER. (SMS)

2.16.2. DOCUMENTATION REFERENCE

KOPLYAY, J.B. REFERENCE MANUAL FOR IAP: COMPUTER BASED ITEM ANALYSIS PROGRAM FOR ANALYZING ACHIEVEMENT TYPE TESTS AND/OR ITEMS. UNPUBLISHED MANUSCRIPT, 1974.

2.17. IAPG (ANALYSIS OF QUESTIONNAIRE TYPE TESTS)

2.17.1. ABSTRACT

IAPG IS A VERSATILE SERIES OF ITEM ANALYSIS COMPUTER PROGRAMS WHICH ARE TO BE USED AS FIELD PREDICTION INSTRUMENTS. IAPG PRINTED RESULTS ENABLE THE USER TO CREATE A COMPOSITE (SCALE) WHICH OPTIMALLY PREDICTS A GIVEN CRITERION AND WHICH CAN BE QUICKLY AND EASILY SCORED BY HAND. TO FACILITATE HAND SCORING, THE COMPOSITE SCORE, I.E., THE WEIGHTED SUM OF SCORES FOR THE QUESTIONNAIRE ITEMS, IS COMPUTED USING ONLY UNIT (+1 OR -1) WEIGHTS FOR EACH ITEM. THE INPUT CONSISTS OF RESPONSES TO ITEMS, SUCH AS BIOGRAPHICAL, INTEREST, OR OPINION-TYPE QUESTIONS. THE DATA SET OF RESPONSES MUST BE DIVIDED INTO TWO OR THREE SUBSAMPLES. THE SELECTION OF MORE THAN ONE ALTERNATIVE OF ANY ITEM IS NOT ALLOWED. IF DESIRED, OMITTED OR OUT-OF-RANGE RESPONSES CAN BE CONSIDERED AS AN ALTERNATIVE; OTHERWISE, ALL CASES WITH OMITTED OR OUT-OF-RANGE RESPONSES ARE ELIMINATED. IAPG IS DIVIDED INTO TWO LOGICAL UNITS, IAPG/1-4 AND IAPG/5-7. IAPG/1-4 COMPUTES ITEM KEYS (A KEY FOR AN ITEM WITH K ALTERNATIVES IS AN ORDERED SET OF K NUMBERS SUCH THAT ALL INDIVIDUALS RESPONDING TO THE I-TH ALTERNATIVE RECEIVE THE I-TH NUMBER AS THEIR (KEY) SCORE FOR THAT ITEM) AND ITEM VALIDITIES ON EACH SUBSAMPLE, AND CROSS-VALIDATES THE RESULTS ON THE REMAINING SUBSAMPLES. IAPG/5-7 USES ITEMS SELECTED ON THE BASIS OF IAPG/1-4 OUTPUT TO FORM THE BEST UNIT-WEIGHTED PREDICTOR OF EACH SPECIFIED CRITERION ON EACH SUBSAMPLE AND CROSS-VALIDATES THE RESULTS ON THE REMAINING SUBSAMPLES. (SMS)

2.17.2. DOCUMENTATION REFERENCE

ALBERT, W.G., & WHITEHEAD, L.K. IAPG: AN ITEM ANALYSIS PROGRAM FOR QUESTIONNAIRE TYPE TEST INSTRUMENTS. UNPUBLISHED MANUSCRIPT, 1975.

2.18. IAPG-HITS (HIT-TABLE FOR DICHOTOMOUS CRITERIA)

2.18.1. ABSTRACT

IAPG-HITS IS A UTILITY PROGRAM USED TO SUPPLEMENT THE IAPG/5-7 "ITEM ANALYSIS PROGRAM FOR QUESTIONNAIRE DATA" BY DETERMINING THE ACCURACY OF PREDICTION OF DICHOTOMOUS OR ARTIFICIALLY DICHOTOMIZED CRITERIA. THIS IS ACCOMPLISHED BY COMPUTING BINARY CLASSIFICATION TABLES FOR THE DICHOTOMOUS (0/1) CRITERIA USING SPECIAL VERSIONS OF THE PERSUB TABDEV AND HITS SUBROUTINES (SEE SECTION 3.3 FOR A DESCRIPTION OF PERSUB, TABDEV, AND HITS). THE INPUT FILES ARE THE "ISSP" (ITEM SELECTION SEQUENCE FILE) AND THE "KIRF" (KEYED ITEM RESPONSE FILE) GENERATED BY THE IAPG/5-7 PROGRAM. THE CRITERION MAY BE EITHER DICHOTOMOUS OR CONTINUOUS. HOWEVER, THE CONTINUOUS CRITERION WILL BE DICHOTOMIZED ACCORDING TO A USER SUPPLIED CUTOFF VALUE. OTHER OPTIONS INCLUDE A FILE CHARACTERISTICS LIST AND A PRINT OF THE TABLE DEVELOPMENT FOR THE 100 CUTOFF POINTS BETWEEN 1 AND 0.

THE NUMBER OF CASES IS LIMITED TO 2000 PER SAMPLE, WHILE THE NUMBER OF ITEMS INCLUDING DUMMIES, MUST BE 400 OR LESS.
(SMS)

2.18.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - IAPG-HITS.

IAPG-HITS REPORT RETRIEVAL COMMAND: @T.T.DOC IAPG-HITS.

2.19. MANOVA (MULTIVARIATE ANALYSIS OF VARIANCE)

2.19.1. ABSTRACT

MANOVA IS COMPOSED OF TWO PROGRAMS,"BMDX69X" AND "X69HT," WHICH TOGETHER WILL SOLVE A VARIETY OF MULTIVARIATE ANALYSIS OF VARIANCE AND COVARIANCE PROBLEMS.

BMDX69X:
THIS PROGRAM, WHICH IS A MODIFICATION OF THE BMD PROGRAM BMDX69, WILL PERFORM MODEL 1 MULTIVARIATE ANALYSIS OF VARIANCE OF VARIANCE OR COVARIANCE FOR ANY HIERARCHICAL DESIGN WITH EQUAL CELL SIZES. THIS INCLUDES NESTED, PARTIALLY NESTED AND CROSSED, AND FULLY CROSSED DESIGNS. THE DESIGN IS SPECIFIED BY INDICATING THE NESTING RELATIONSHIPS OF THE INDICES. SEVERAL ANALYSES MAY BE PERFORMED FOR EACH PROBLEM BY SPECIFYING DIFFERENT DEPENDENT VARIABLES OR COVARIATES.

X69HT:
THIS PROGRAM WILL TAKE THE COVARIANCE MATRICES FILE GENERATED BY BMDX69X AND COMPUTE A NEW MANOVA ANALYSIS WITH AN ERROR MATRIX GENERATED FROM SPECIFIED ERROR COMPONENTS. UNIVARIATE TESTS MAY BE PERFORMED FOR EACH DEPENDENT VARIABLE.

THE MANOVA PACKAGES ARE LIMITED TO 10 ANALYSIS OF VARIANCE INDICES (FACTORS) AND 40 DEPENDENT VARIABLES. CORE USED BY THE BMDX69X PACKAGE IS VARIABLE AND IS ACQUIRED AS NEEDED BY THE PROGRAM. (SMS)

2.19.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - MANOVA.

MANOVA REPORT RETRIEVAL COMMAND: @T*T.DOC MANOVA.

2.20. MAX-FACTOR (FACTOR ANALYSIS)

2.20.1. ABSTRACT

MAX-FACTOR IS A FACTOR ANALYSIS PROGRAM FOR PROBLEMS HAVING UP TO 400 VARIABLES. THE POWER METHOD IS USED TO COMPUTE THE FACTORS ONE AT A TIME IN DESCENDING ORDER ACCORDING TO THE PERCENTAGE OF THE TOTAL VARIANCE EACH FACTOR EXPLAINS. THIS FEATURE ENABLES THE USER TO COMPUTE ONLY AS MANY FACTORS AS ARE REQUIRED. ADDITIONAL FACTORS MAY BE CALCULATED BY RESTARTING THE PROGRAM AT THE PRIOR TERMINATION POINT. THE INPUT CONSISTS OF THE UPPER TRIANGULAR PORTION OF A CORRELATION MATRIX. THE DIAGONAL ENTRIES MAY BE ONES, IN WHICH CASE A PRINCIPAL COMPONENTS ANALYSIS IS PERFORMED, OR THE DIAGONAL ENTRIES MAY BE USER SPECIFIED ESTIMATES OF THE COMMUNALITIES (SUCH AS THE LARGEST CORRELATION IN EACH ROW OF THE MATRIX, OR THE SQUARED MULTIPLE CORRELATION OBTAINED FROM THE REGRESSION OF EACH VARIABLE ON THE REMAINING VARIABLES), OR THE DIAGONAL ENTRIES MAY BE READ FROM A FILE. THE USER MAY REQUEST THAT THE PROGRAM BE USED ITERATIVELY TO IMPROVE INITIAL ESTIMATES OF THE COMMUNALITIES. THE FACTOR LOADINGS CAN BE SAVED FOR SUBSEQUENT USE. ANY SUBSET OF THE FACTORS MAY BE ROTATED TO A SIMPLE STRUCTURE FORM USING THE VARIMAX CRITERION. IF DESIRED, THE FACTOR LOADINGS CAN BE SORTED ACCORDING TO ABSOLUTE VALUE. (SMS)

2.20.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - MAX-FACTOR.

MAX-FACTOR REPORT RETRIEVAL COMMAND: @T.T.DOC MAX-FACTOR.

2.21. MEANSD (MEAN & STANDARD DEVIATION PROGRAM)

2.21.1. ABSTRACT

THE PURPOSE OF THIS PROGRAM IS TO ENABLE THE USER TO CALCULATE THE MEAN AND STANDARD DEVIATION OF THE MINOR VALUES FOR EACH CHANGE IN THE MAJOR VALUE ON A DIG FREQUENCY FILE.

ALL CALCULATIONS ARE PERFORMED IN DOUBLE PRECISION ARITHMETIC IN ORDER TO RETAIN THE MAXIMUM ACCURACY IN THE RESULTS.

EACH DISTRIBUTION ON THE INPUT FILE TO BE EXTRACTED FOR COMPUTATIONS IS IDENTIFIED BY A SEPARATE CONTROL CARD. AN OUTPUT RECORD FOR EACH CHANGE IN THE MAJOR FIELD WITHIN EACH DISTRIBUTION IS WRITTEN. EACH OUTPUT RECORD CONTAINS

A. THE DISTRIBUTION IDENTIFIER,
B. MAJOR FIELD VALUE,
C. MEAN,
D. STANDARD DEVIATION,
E. VALID-N (CASES WITH ALL NUMERIC DATA),
F. TOTAL-N (ALL CASES),
G. SUM OF X, AND
H. SUM OF X-SQUARED.

THE OUTPUT FILE MAY BE USED AS INPUT TO ADDITIONAL STATISTICAL ANALYSES OR REPORTED BY STANDARD REPORT WRITING PROGRAMS SUCH AS VIRUS. (SMA)

2.21.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - SMAZ UTILITY PROGRAMS & PROCESSORS.

REPORT RETRIEVAL COMMAND: @Z.Z.DOC.

2.22. MULT-DISCR (MULTIPLE DISCRIMINANT ANALYSIS)

2.22.1. ABSTRACT

MULTI-DISCR IS A MAJOR MODIFICATION OF DISCRM FROM D.J. VELDMAN'S "FORTRAN PROGRAMMING FOR THE BEHAVIORAL SCIENCES", AND WILL PERFORM MULTIPLE DISCRIMINANT ANALYSIS FOR UP TO TEN (AND IN SOME CASES MORE) GROUPS HAVING UP TO 100 VARIABLES. THE DYNAMIC ALLOCATION OF CORE ALLOWS FOR MORE THAN TEN GROUPS AND 100 VARIABLES. THE PROGRAM HAS AN OPTION TO COMPUTE THE AMOUNT OF CORE REQUIRED FOR A PARTICULAR PROBLEM SO THE USER MAY DETERMINE WHETHER HIS PROBLEM MAY BE RUN. A PROBLEM WITH 149 VARIABLES AND 4 GROUPS, READING A COBOL FILE, REQUIRED 93K TO RUN.

OUTPUT FROM THE PROGRAM INCLUDES A SERIES OF LINEAR COMBINATIONS OF THE VARIABLES CALLED THE DISCRIMINANT FUNCTIONS, AS WELL AS A TEST OF SIGNIFICANCE OF EACH FUNCTION; GROUP AND TOTAL SAMPLE MEANS, STANDARD DEVIATIONS, AND CORRELATION MATRICES FOR THE ORIGINAL VARIABLES; CORRELATIONS BETWEEN THE ORIGINAL VARIABLES AND THE DISCRIMINANT FUNCTIONS; AND THE MEANS OF THE SCORES WITHIN EACH GROUP (CENTROIDS).

A NEW FEATURE OF THE PROGRAM IS THE ABILITY TO COMPUTE THE PROBABILITY THAT AN INDIVIDUAL BELONGS TO ONE OF THE GROUPS GIVEN HIS VARIABLE SCORES. FOR EACH GROUP, A TABLE IS PRINTED SHOWING HOW THE PROBABILITIES ARE DISTRIBUTED WITHIN THAT GROUP. (SMS)

2.22.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - MULT-DISCR.

MULT-DISCR REPORT RETRIEVAL COMMAND: @T.T.DOC MULT-DISCR.

2.23. MULTI-PROB (MULTIVARIATE NORMAL PROBABILITY)

2.23.1. ABSTRACT

PERSONNEL RESEARCH IS FREQUENTLY CONFRONTED BY PROBLEMS WHICH INVOLVE INVESTIGATION OF THE JOINT OCCURRENCE OF SEVERAL ATTRIBUTES OR TRAITS WHICH ARE NORMALLY DISTRIBUTED IN A POPULATION. FREQUENTLY, THE PROBLEM IS ONE OF DETERMINING THE PROPORTION OF A POPULATION WHICH WOULD BE EXPECTED TO POSSESS THESE MEASURES: E.G., APTITUDE INDICES OR THE LIKE, COMMONLY WITHIN SPECIFIED RANGES (CUTOFFS). THESE PROPORTIONS ARE REPRESENTED BY A SINGLE NUMBER WHICH IS USUALLY CALLED A MULTIVARIATE PROBABILITY.

THE PROBLEM OF COMPUTING THESE PROBABILITIES WITH SATISFACTORY PRECISION BEYOND THE TWO-VARIATE CASE HAS BEEN, IN GENERAL, AN UNSOLVED ONE. PREVIOUSLY PUBLISHED PAPERS BY OWEN (1956) AND STECK (1958) GAVE METHODS APPLICABLE TO THE TWO- AND THREE-VARIABLE CASE. THEIR METHODS WERE MODIFIED AND EXTENDED TO FOUR AND FIVE VARIABLES IN ORDER TO PRODUCE THE SYSTEM OF DIRECT CALCULATION USED BY MULTI-PROB.

THE MULTI-PROB PROGRAM IS DESIGNED TO COMPUTE MULTIVARIATE NORMAL PROBABILITIES FOR UP TO FIVE JOINTLY DISTRIBUTED NORMAL VARIABLES EACH HAVING MEAN 0 AND VARIANCE OF 1. GIVEN THE PROPORTIONS U(1),U(2),...,U(N), AND THE CORRELATION MATRIX FOR A SET OF RANDOM VARIABLES X(1),X(2),...,X(N), EACH WITH MEAN 0 AND VARIANCE 1 WHOSE JOINT DISTRIBUTION IS MULTIVARIATE NORMAL, THE PROGRAM COMPUTES THE PROBABILITY P(X(1).LE.U(1),...,X(N).LE.U(N)). THE NUMBER OF VARIABLES, THE CORRELATION MATRIX, AND THE CUTOFF PROPORTIONS ARE OBTAINED BY CARD INPUT.

THE APPENDIX OF THE MULTI-PROB DOC PROCESSOR REPORT WILL BE OF INTEREST TO THOSE WHO WISH TO COMPUTE MULTI-VARIATE PROBABILITIES AS PART OF THEIR OWN PROGRAMS. (SMS)

2.23.2. DOCUMENTATION REFERENCES

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE - MULTI-PROB.

MULTI-PROB REPORT RETRIEVAL COMMAND: @RT.DOC MULTI-PROB.

2.29. REGREX (STEPWISE REGRESSION)

2.29.1. ABSTRACT

REGREX IS A COMPUTER-BASED STEPWISE REGRESSION ALGORITHM, DEVELOPED IN-HOUSE, WHICH HANDLES UP TO 200 VARIABLES. IT GREW OUT OF THE NEED OF RESEARCH SCIENTISTS FOR A MORE VERSATILE, HIGHLY ACCURATE, AND FLEXIBLE PREDICTION SYSTEM TO SUPPLEMENT THE EXISTING ITERATIVE ALGORITHM. FIRST, REGREX WILL AUTOMATICALLY IDENTIFY REDUNDANT VARIABLES, ELIMINATE THEM FROM THE REGRESSION, AND REPORT THE PROPER DEGREES OF FREEDOM FOR F-TESTS. SECOND, REGREX ALLOWS FOR EXTENSIVE RESIDUAL ANALYSIS INCLUDING RESIDUAL PLOTS, TO EXAMINE THE FIT OF NORMALITY AND HOMOSCEDASTICITY. THIRD, REGREX PROVIDES APPROPRIATE REGRESSION WEIGHTS AT EACH STEP. FOURTH, THE PRINTED OUTPUT IS MORE DETAILED, ALLOWING MORE PRECISE AND PROPER INTERPRETATION OF THE RESULTS. FIFTH, SEVERAL OPTIONS ARE INCORPORATED INTO REGREX WHICH ALLOW GREAT FLEXIBILITY FOR BOTH RESEARCH SCIENTISTS AND AIR FORCE MANAGERS. ONE OF THESE OPTIONS ALLOWS NEW VARIABLES TO BE GENERATED FROM COMBINATIONS OF EXISTING VARIABLES. ANOTHER OPTION ALLOWS PREDICTORS TO BE "FORCED" INTO THE PREDICTION SYSTEM REGARDLESS OF THEIR RELATIVE OR ABSOLUTE CONTRIBUTION IN EXPLAINING THE CRITERION VARIANCE. (SMS)

2.29.2. DOCUMENTATION REFERENCES

KOPLYAY, J.B., MATHON, W.S., & WHITEHEAD, L.R. REGREX: A COMPUTER BASED STEPWISE REGRESSION ALGORITHM WITH RESIDUAL ANALYSIS. UNPUBLISHED MANUSCRIPT, 1976.

AFMRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE = REGREX.

REGREX REPORT RETRIEVAL COMMAND: WT*T.DOC REGREX.

8.25. TRICOR (CORRELATION AND REGRESSION PROGRAM)

8.25.1. ABSTRACT

TRICOR IS A LARGE MODULAR PROGRAM WHICH PERFORMS A VARIETY OF STATISTICAL ANALYSIS PROCEDURES FROM A STANDARD COBOL INPUT FILE. THE OPERATIONS TO BE PERFORMED ARE DEFINED BY THE PROGRAMMER VIA A SET OF FREE-FORMAT COMMANDS. TRICOR WILL PERFORM A RIGOROUS ANALYSIS OF THESE COMMANDS, AND IF NO ERRORS ARE DETECTED, WILL PERFORM THE DESIGNATED OPERATIONS.

THE BASIC OPERATIONS PERFORMED BY TRICOR ARE

A. CONVERT RAW DATA FROM COBOL FILES TO FORTRAN FILES WITH A MAXIMUM OF 30,000 CASES IN UP TO 500 SAMPLES.
B. COMPUTE NEW VARIABLES USING THOSE WHICH WERE CONVERTED FROM THE INPUT WITH A MAXIMUM OF 400 VARIABLES.
C. COMPUTE A CORRELATION MATRIX WITH THE OPTION OF SAVING IT ON A MASTER FILE.
D. COMPUTE STEPWISE AND/OR ITERATIVE REGRESSION ANALYSIS WITH THE OPTION OF SAVING THE REGRESSION WEIGHTS.
E. COMPUTE F-STATISTICS, AND
F. COMPUTE PREDICTED SCORES FOR CROSS-VALIDATION.

ALL PRINTED REPORTS PRODUCED BY TRICOR CONFORM TO THE AFHRL/SM STANDARDS FOR FINAL REPORT PRODUCTS.

INTERMEDIATE WORK FILES GENERATED BY TRICOR MAY BE SAVED AND USED WITH OTHER UTILITY PROGRAMS OR SPECIAL-PURPOSE PROGRAMS DEVELOPED BY APPLICATIONS PROGRAMMERS. (SMA)

8.25.2. DOCUMENTATION REFERENCE

AFHRL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE = UTILITY CORRELATION & REGRESSION SYSTEM.

REPORT RETRIEVAL COMMAND: @Z*8.DOC.

2.26. APRIME (RESIDUALS TEST FOR NORMALITY)

2.26.1. ABSTRACT

APRIME IS A COMPUTER PROGRAM WHICH PERFORMS AN APPROXIMATE ANALYSIS OF VARIANCE TEST FOR NORMALITY ON THE RESIDUALS FROM A REGRESSION PROBLEM. THE MATHEMATICAL THEORY UNDERLYING THE APRIME TEST FOR NORMALITY IS DOCUMENTED IN SHAPIRO AND FRANCIA (1972). MONTE CARLO RESULTS MADE SHOWN THAT FOR MANY ALTERNATIVE DISTRIBUTIONS THE APRIME TEST HAS POWER AS GOOD OR BETTER THAN THE FOLLOWING TEST PROCEDURES: SKEWNESS, KURTOSIS, CHI-SQUARE, DURBIN, KOLMOGOROV-SMIRNOV, CRAMER-VON MISES, WEIGHTED CRAMER-VON MISES, AND DAVID'S U. THE PRINCIPAL DRAWBACK TO THE USE OF APRIME IS THAT PERCENTAGE POINTS ARE NOT AVAILABLE FOR SAMPLES OF SIZE GREATER THAN 99.

THE INPUT FILE FOR APRIME IS THE OUTPUT FROM THE TOICOR PREDICTED SCORE LINE. IF THE NUMBER OF CASES ON THE OUTPUT FILE EXCEEDS 99, A PRELIMINARY RUN MUST BE MADE TO SELECT A RANDOM SUBSAMPLE OF SIZE 99. (SMS)

2.26.2. DOCUMENTATION REFERENCE

AFGWL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE = APRIME.

APRIME REPORT RETRIEVAL COMMAND: @T=T.DOC APRIME.

8.27. USER SUPPLIED FREQUENCY DISTRIBUTION PLOTS

8.27.1. ABSTRACT

This program prints user-supplied frequency distributions generated by SIGA. Three format control commands direct the way in which the program will

1. Select the distributions to be printed.
2. Designate values to be "pulled" or plotted.
3. Compute mean, standard deviation, and median.
4. Designate the range of values to be used for computations.
5. Provide report title information, and
6. Name a utility file layout for distribution requests.

User reports are especially suited for the display of survey data.

The printed reports produced by user conform to the group/on standards for final report products. (SRI)

8.27.2. DOCUMENTATION REFERENCE

Group Unique User Documentation Report (unpublished): Title = SAS utility programs & procedures.

Report retrieval command: QIC-DOC.

3.1. UNIVAC MATH-PACK

3.1.1. GENERAL DESCRIPTION

THE UNIVAC MATH-PACK SYSTEM CONTAINS A COMPREHENSIVE SET OF 70 MATHEMATICAL SUBROUTINES DESIGNED FOR UNIVAC LARGE SCALE SYSTEMS. THE SYSTEM PROVIDES THE MORE FREQUENTLY USED TECHNIQUES IN NUMERICAL ANALYSIS, WITH EACH SUBROUTINE DESIGNED SO THAT THE CAPABILITIES OF THE UNIVAC LARGE SCALE EQUIPMENT ARE USED MOST EFFICIENTLY (E.G., WITH RESPECT TO STORAGE REQUIREMENTS, INTERNAL COMPUTER SPEED, AND ACCURACY), AND THE PROGRAM PREPARATION REQUIRED PRIOR TO CALLING THE SUBROUTINE IS MINIMIZED. THE MATH-PACK SUBROUTINES ARE CATEGORIZED INTO 13 GROUPS AND ARE LISTED IN THE NEXT SECTION.

3.1.2. LIST OF SUBROUTINES

INTERPOLATION

GNINT - GREGORY-NEWTON INTERPOLATION
GNEXT - GREGORY-NEWTON EXTRAPOLATION
GNPOL - GREGORY-NEWTON POLYNOMIAL EVALUATION
BESINT - BESSEL INTERPOLATION
STINT - STIRLING INTERPOLATION
CDINT - GAUSS CENTRAL-DIFFERENCE INTERPOLATION
AITINT - AITKEN INTERPOLATION
YLGINT - LAGRANGE INTERPOLATION
SPLN1, SPLN2 - SPLINE INTERPOLATION

NUMERICAL INTEGRATION

TRAPN1 - TRAPEZOIDAL RULE
SIMIN1 - SIMPSON 1/3 RULE
SIM3N1 - SIMPSON 3/8 RULE
STEPN1 - VARIABLE STEP INTEGRATION
GENN1 - GENERALIZED NUMERICAL QUADRATURE
DOUBN1 - DOUBLE INTEGRATION
LGAUSS - GAUSS QUADRATURE ABSCISSAS AND WEIGHTS
SIMPTS - SIMPSON 1/3 RULE ABSCISSAS AND WEIGHTS

MATHEMATICAL AND STATISTICAL SOFTWARE INDEX

SOLUTION OF EQUATIONS

NEWTIT - NEWTON-RAPHSON ITERATION
WEGIT - WEGSTEIN ITERATION
AITIT - AITKEN ITERATION
ROOTCP - REAL AND COMPLEX ROOTS OF A REAL OR COMPLEX POLYNOMIAL

DIFFERENTIATION

DERIV1 - FIRST DERIVATIVE APPROXIMATION
DERIV2 - SECOND DERIVATIVE APPROXIMATION
NTHDER - NTH DERIVATIVE OF A POLYNOMIAL

POLYNOMIAL MANIPULATION

GIVZRS - POLYNOMIAL COEFFICIENTS GIVEN ITS ZEROS
CVALUE - COMPLEX POLYNOMIAL EVALUATION
POLYX - REAL POLYNOMIAL MULTIPLICATION
CPOLYX - COMPLEX POLYNOMIAL MULTIPLICATION

MATRIX MANIPULATION: REAL MATRICES

MXADD - MATRIX ADDITION
MXSUB - MATRIX SUBTRACTION
MXTRN - MATRIX TRANSPOSITION
MXSCA - MATRIX MULTIPLICATION BY A SCALAR
MXMLT - MATRIX MULTIPLICATION
MXMDIG - MATRIX MULTIPLICATION BY DIAGONAL MATRIX STORED AS A VECTOR
GJR - DETERMINANT: INVERSE: SOLUTION OF SIMULTANEOUS EQUATIONS
MXMOI - INVERSE ACCURACY IMPROVEMENT

MATRIX MANIPULATION: COMPLEX MATRICES

CMXADD - MATRIX ADDITION
CMXSUB - MATRIX SUBTRACTION
CMXTRN - MATRIX TRANSPOSITION
CMXSCA - MATRIX MULTIPLICATION BY A SCALAR

MATHEMATICAL AND STATISTICAL SOFTWARE INDEX

CMXMLT - MATRIX MULTIPLICATION
CGJR - COMPLEX DETERMINANT; INVERSE; SOLUTION OF SIMULTANEOUS EQUATIONS

MATRIX MANIPULATION: EIGENVALUES AND EIGENVECTORS

TRIDMX - TRIDIAGONALIZATION OF REAL SYMMETRIC MATRIX
EIGVAL - EIGENVALUES OF TRIDIAGONAL MATRIX BY STURM SEQUENCES
EIGVEC - EIGENVECTORS OF TRIDIAGONAL MATRIX BY WILKINSON'S METHOD

MATRIX MANIPULATION: MISCELLANEOUS

DGJR - DOUBLE-PRECISION DETERMINANT; INVERSE; SOLUTION OF SIMULTANEOUS EQUATIONS
PMXTRI - POLYNOMIAL MATRIX TRIANGULARIZATION
SCALE - POLYNOMIAL MATRIX SCALING
MXROT - MATRIX ROTATION

ORDINARY DIFFERENTIAL EQUATIONS

EULDE - EULER'S METHOD
HAMDE - HAMMING'S METHOD
INVAL - INITIAL VALUES FOR DIFFERENTIAL EQUATION SOLUTION
RKDE - RUNGE-KUTTA METHOD
SODE - SECOND-ORDER EQUATIONS
MRKDE - REDUCTION OF MTH-ORDER SYSTEM TO SYSTEM OF M FIRST-ORDER EQUATIONS

SYSTEMS OF EQUATIONS

JACMX - JACOBI ITERATION TO DETERMINE EIGENVALUES AND EIGENVECTORS OF SYMMETRIC MATRIX
HJACMX - JACOBI ITERATIONS TO DETERMINE EIGENVALUES AND EIGENVECTORS OF HERMITIAN MATRIX
LSIMEQ - SOLUTION TO A SET OF LINEAR SIMULTANEOUS EQUATIONS
NSIMEQ - FUNCTIONAL ITERATION TO DETERMINE SOLUTION TO SET OF NON-LINEAR EQUATIONS

MATHEMATICAL AND STATISTICAL SOFTWARE INDEX

CURVE FITTING

CFSRIE - COEFFICIENTS OF FOURIER SERIES ON A CONTINUOUS RANGE
FTRANS - FOURIER TRANSFORM
DFSRIE - COEFFICIENTS OF FOURIER SERIES ON DISCRETE RANGE
FITD - FITTED VALUE AND DERIVATIVE VALUES FOR A LEAST-SQUARES POLYNOMIAL
ORTHLS - ORTHOGONAL POLYNOMIAL LEAST-SQUARES CURVE FITTING
FITY - FITTED VALUES FOR A LEAST-SQUARES POLYNOMIAL
COEFS - COEFFICIENTS OF A LEAST-SQUARES POLYNOMIAL

PSEUDO RANDOM NUMBER GENERATORS

NRAND - INTERVAL (0,2**27) GENERATOR
RANDU - UNIFORM DISTRIBUTION
RANDN - NORMAL DISTRIBUTION
RANDEX - EXPONENTIAL DISTRIBUTION

3.1.3. DOCUMENTATION REFERENCES

UNIVAC SYSTEMS REFERENCE LIBRARY. UNIVAC LARGE SCALE SYSTEMS MATH-PACK PROGRAM ABSTRACTS. UP-4051 REV. 2, SPERRY RAND CORPORATION, 1970.

UNIVAC SYSTEMS REFERENCE LIBRARY. UNIVAC LARGE SCALE SYSTEMS MATH-PACK PROGRAMMERS REFERENCE. UP-7542 REV. 1, SPERRY RAND CORPORATION, 1970.

THE FIRST REFERENCE CONTAINS A SET OF ABSTRACTS SUMMARIZING THE PURPOSE, CALLING SEQUENCE, AND METHOD OF COMPUTATION FOR EACH MATH-PACK SUBROUTINE. THE SECOND REFERENCE GIVES MORE DETAILED INFORMATION CONCERNING EACH SUBROUTINE.

3.2. UNIVAC STAT-PACK

3.2.1. GENERAL DESCRIPTION

THE UNIVAC STAT-PACK SYSTEM IS COMPRISED OF 78 FUNDAMENTAL STATISTICAL SUBROUTINES INVOLVING THE MORE FREQUENTLY USED STATISTICAL TECHNIQUES. LIKE THE SUBROUTINES IN THE UNIVAC MATH-PACK SYSTEM, THE STAT-PACK SUBROUTINES ARE DESIGNED TO MINIMIZE THE PROGRAM PREPARATION REQUIRED PRIOR TO CALLING A SUBPROGRAM AND TO OPTIMIZE ON A COMBINATION OF DESIRABLE FEATURES INCLUDING INTERNAL COMPUTER SPEED, ACCURACY, AND STORAGE REQUIREMENTS. THE STAT-PACK SUBROUTINES ARE GROUPED INTO 11 GENERAL STATISTICAL CATEGORIES AND ARE LISTED IN THE FOLLOWING SECTION.

3.2.2. LIST OF SUBROUTINES

DESCRIPTIVE STATISTICS

FREQP - FREQUENCY POLYGON
HIST - HISTOGRAM
MHIST - MULTIVARIATE HISTOGRAM
GROUP - GROUPING OF DATA

ELEMENTARY POPULATION STATISTICS

AMEAN - ARITHMETIC MEAN
GMEAN - GEOMETRIC MEAN
HMEAN - HARMONIC MEAN
MEDIAN - MEDIAN
MODE - MODE
QUANT - QUANTILES
OGIVE - DISTRIBUTION CURVE
IQRNG - INTERPERCENTILE RANGE
RANGE - RANGE
MNDEV - MEAN DEVIATION
STDEV - STANDARD DEVIATION
CVAR - COEFFICIENT OF VARIATION
ORDER - ORDER AND RANK STATISTICS

MATHEMATICAL AND STATISTICAL SOFTWARE INDEX

CMONT - CENTRAL MOMENTS
AMONT - ABSOLUTE MOMENTS
CUMLT - CUMULANTS
SHPCOR - SHEPPARD'S CORRECTIONS
KURSK - SKEWNESS AND KURTOSIS

DISTRIBUTION, FITTING AND PLOTTING

BINOM - BINOMIAL DISTRIBUTION
POISON - POISSON DISTRIBUTION
HYPER - HYPERGEOMETRIC DISTRIBUTION
PNORM - NORMAL DISTRIBUTION
AFSER - ARNE FISHER SERIES

CHI-SQUARE TESTS

CHI2IS - CHI-SQUARE TEST OF SAMPLE PROPORTION FOR ONE SAMPLE
CHI2JS - CHI-SQUARE TEST OF SAMPLE PROPORTION FOR J SAMPLES
CHI2P - CHI-SQUARE TEST OF FIT TO POISSON DISTRIBUTION
CHI2N - CHI-SQUARE TEST OF NORMALITY
CHISAM - CHI-SQUARE TEST OF HOMOGENEITY
CHICNT - CHI-SQUARE TEST FOR INDEPENDENCE
GENGOF - CHI-SQUARE TEST OF GENERAL GOODNESS OF FIT

SIGNIFICANCE TESTS

SIGPRP - TEST OF SIGNIFICANCE OF PROPORTION OF SUCCESSES
SIGMN - TEST OF SIGNIFICANCE OF A MEAN
SIGDMN - TEST OF SIGNIFICANCE OF THE DIFFERENCE BETWEEN TWO MEANS
SIGDVR - TEST OF SIGNIFICNCE OF THE RATIO BETWEEN TWO VARIANCES

CONFIDENCE INTERVALS

CFDMKV - CONFIDENCE INTERVAL FOR THE MEAN: KNOWN VARIANCE
CFDMUV - CONFIDENCE INTERVAL FOR THE MEAN: UNKNOWN VARIANCE
CFDMSU - CONFIDENCE INTERVAL FOR THE DIFFERENCE BETWEEN TWO MEANS

MATHEMATICAL AND STATISTICAL SOFTWARE INDEX

CFDVAR - CONFIDENCE INTERVAL FOR VARIANCE
TOLINT - TOLERANCE INTERVALS

ANALYSIS OF VARIANCE

ANOV1 - ONE-WAY CROSS CLASSIFICATION
ANOV2 - TWO-WAY CROSS CLASSIFICATION
ANOV3 - THREE-WAY CROSS CLASSIFICATION
MISDAT - MISSING DATA
VTRANS - VARIABLE TRANSFORMATIONS
ANOVRB - RANDOMIZED BLOCKS
ANOVLS - LATIN SQUARES
ANOVSP - SPLIT-PLOT DESIGN
ANOSSP - SPLIT-SPLIT PLOT DESIGN
ANOVN2 - TWO-WAY NESTED DESIGN
ANOVN3 - THREE-WAY NESTED DESIGN
ANOCO - ANALYSIS OF COVARIANCE
GLH - GENERAL LINEAR HYPOTHESES

REGRESSION ANALYSIS

RESTEM - STEPWISE MULTIPLE REGRESSION
REBSOM - BACK SOLUTION MULTIPLE REGRESSION
CORAN - CORRELATION ANALYSIS

TIME SERIES ANALYSIS

MOVAVG - MOVING AVERAGES
SEASHI - SHISKIN'S SEASONALITY FACTORS
WEMAV - WEIGHTED MOVING AVERAGES
TRELS - TREND ANALYSIS BY LEAST SQUARES
VADIME - VARIATE DIFFERENCE METHOD
TSFARG - AUTOREGRESSIVE MODEL
GEXSMO - GENERALIZED EXPONENTIAL SMOOTHING
AUXCOR - AUTO-CORRELATION AND CROSS-CORRELATION ANALYSIS
POWDEN - POWER DENSITY FUNCTIONS
RCPROB - RESIDUAL PROBABILITIES

MULTI-VARIATE ANALYSIS

GENVAR - GENERALIZED VARIANCE
DISHOT - HOTELLING'S DISTRIBUTION

DSQ - MAHALANOBIS' DISTRIBUTION
SIGIMN - SIGNIFICANCE OF A SET OF MEANS
DISCRA - DISCRIMINANT ANALYSIS
FACTAN - FACTOR AND PRINCIPAL COMPONENT ANALYSIS

MISCELLANEOUS SUBROUTINES

PLOTI - PLOT OF ONE LINE
JIM - MATRIX INVERSION
MXTMLT - LEFT MULTIPLICATION OF A MATRIX BY ITS TRANSPOSE

3.2.3. DOCUMENTATION REFERENCES

UNIVAC SYSTEMS REFERENCE LIBRARY. UNIVAC 1106 SYSTEM/1108 MULTI-PROCESSOR SYSTEM STAT-PACK PROGRAM ABSTRACTS. UP-4041 REV. 2, SPERRY RAND CORPORATION, 1969.

UNIVAC SYSTEMS REFERENCE LIBRARY. UNIVAC LARGE SCALE SYSTEMS STAT-PACK PROGRAMMERS REFERENCE. UP-7502 REV. 1, SPERRY RAND CORPORATION, 1970.

THE FIRST REFERENCE CONTAINS A SET OF ABSTRACTS SUMMARIZING THE PURPOSE, CALLING SEQUENCE, AND METHOD OF COMPUTATION FOR THE STAT-PACK SUBROUTINES. THE SECOND REFERENCE GIVES MORE DETAILED INFORMATION CONCERNING EACH SUBROUTINE.

3.3. PERSUB SYSTEM

3.3.1. GENERAL DESCRIPTION

THE PERSUB SUBROUTINE SYSTEM IS A SET OF MATRIX-ORIENTED SUBROUTINES DEVELOPED DURING THE MID-SIXTIES AT THE PERSONNEL RESEARCH LABORATORY (NOW A PART OF THE AIR FORCE HUMAN RESOURCES LABORATORY), AEROSPACE MEDICAL DIVISION, LACKLAND AIR FORCE BASE, TEXAS. THE SYSTEM EVOLVED FROM WORK IN A PARTICULAR TYPE OF DATA ANALYSIS. MANY OF THE SUBROUTINES IN PERSUB INCORPORATE CLASSICAL STATISTICAL PROCESSES WHICH ARE USED IN PERSONNEL RESEARCH (E.G., CORRELATION, REGRESSION, FACTOR ANALYSIS, MATRIX MANIPULATION, ETC.), AND THE SYSTEM WAS DEVELOPED PRIMARILY FOR THE PURPOSE OF PROVIDING THE RESEARCHER WITH A MAXIMUM OF FLEXIBILITY IN DESIGNING A SEQUENCE OF ANALYSES TO BE CARRIED OUT ON RESEARCH DATA. THE ORIGINAL PERSUB SYSTEM, ALONG WITH OTHER SUBROUTINES WHICH WERE DEVELOPED LATER, CONSISTS OF 69 SUBROUTINES. HOWEVER, ONLY 36 SUBROUTINES, THOSE WHICH HAVE A MATHEMATICAL OR STATISTICAL FUNCTION, WILL BE LISTED BELOW. FOR CONVENIENCE, THE SUBROUTINES HAVE BEEN CATEGORIZED INTO SEVEN GROUPS ACCORDING TO THE GENERAL MATHEMATICAL OR STATISTICAL TECHNIQUE TO WHICH THEY RELATE.

3.3.2. LIST OF SUBROUTINES

DESCRIPTION AND TABULATION

GRAPH - TWO-DIMENSIONAL GRAPH

MATRIX ALGEBRA/MANIPULATION

DETERM - DETERMINANT OF A SYMMETRIC POSITIVE DEFINITE MATRIX
EIGEN - EIGENROOT AND EIGENVECTOR COMPUTATION FOR A SYMMETRIC MATRIX
INVERS - ITERATIVE MATRIX INVERSE
MATADD - MATRIX ADDITION
MATMUL - MATRIX MULTIPLICATION
MATPOW - MATRIX POWERING

MATHEMATICAL AND STATISTICAL SOFTWARE INDEX

MATTRA - MATRIX TRANSPOSE
PTRANS - MATRIX TRANSFORMATION TO PRODUCE AN ORTHOGONAL BASIS (P-TRANSFORMATION)
WEIGHT - WEIGHT ELEMENTS OF A MATRIX
XTRANX - PREMULTIPLICATION OF A MATRIX BY ITS TRANSPOSE

SIGNIFICANCE TESTS/PROBABILITY/DISTRIBUTIONS

BITRUN - MAXIMUM LIKELIHOOD ESTIMATION FROM A TRUNCATED SAMPLE OF THE MEAN AND STANDARD DEVIATION OF A BIVARIATE NORMAL POPULATION
FSTAT - F VALUE ASSOCIATED WITH GIVEN DEGREES OF FREEDOM AND PROBABILITY
MUTRUN - MAXIMUM LIKELIHOOD ESTIMATION FROM A TRUNCATED SAMPLE OF THE MEAN AND STANDARD DEVIATION OF A TRIVARIATE, TETRAVARIATE OR PENTAVARIATE NORMAL POPULATION
PLEVEL - PROBABILITY LEVEL ASSOCIATED WITH GIVEN F STATISTIC AND DEGREES OF FREEDOM
RANDCS - RANDOM DEVIATE GENERATION: CHI-SQUARE DISTRIBUTION
RANEXP - RANDOM DEVIATE GENERATION: EXPONENTIAL DISTRIBUTION
RBTAFT - RANDOM DEVIATE GENERATION: BETA, F OR T DISTRIBUTION
RGAMMA - RANDOM DEVIATE GENERATION: GAMMA DISTRIBUTION
SECANT - MAXIMUM LIKELIHOOD ESTIMATION FROM A TRUNCATED SAMPLE OF THE MEAN AND STANDARD DEVIATION OF A UNIVARIATE NORMAL POPULATION

VARIANCE ANALYSIS

HOMVAR - BOX-SCHEFFE HOMOGENEITY OF VARIANCE TEST

CORRELATION AND REGRESSION ANALYSIS

AKRACY - PREDICTED SCORE ACCURACY DETERMINATION
COMPAR - PREDICTED SCORE AND ACTUAL SCORE COMPARISON
CORRLB - CORRELATION MATRIX
CORRLD - CORRELATION MATRIX (DOUBLE PRECISION)
HITS - 2 X 2 HIT TABLE (USEFUL IN EVALUATING A REGRESSION MODEL WITH A DICHOTOMOUS CRITERION)
PREDSC - PREDICTED SCORES COMPUTATION
REGRED - ITERATIVE REGRESSION

MATHEMATICAL AND STATISTICAL SOFTWARE INDEX

REGRES - ITERATIVE REGRESSION (USEFUL WHEN HIGHLY CORRELATED VARIABLES ARE PRESENT)
TABDEV - HIT TABLE DEVELOPMENT (USEFUL FOR DEVELOPING THE PREDICTION SYSTEM FOR A DICHOTOMOUS CRITERION BY IDENTIFYING THE PREDICTED SCORE CUTTING VALUE WHICH PRODUCES THE SMALLEST NUMBER OF MISCLASSIFICATIONS)
TCFFNT - MATRIX OF T COEFFICIENTS (FOR A GIVEN SET OF VARIABLES, THE I-TH COLUMN CONTAINS THE REGRESSION COEFFICIENTS FOR PREDICTING VARIABLE I FROM THE FIRST (I-1) VARIABLES)

GROUPING OR CLUSTERING

GROUPS - GROUP OBJECTS ON THE BASIS OF AN INPUT MATRIX WHICH CONTAINS MEASURES OF SIMILARITY OR DIFFERENCE BETWEEN ALL PAIRS OF OBJECTS

FACTOR ANALYSIS

COMMUN - COMMUNALITIES COMPUTATION FOR A FACTOR LOADINGS MATRIX
PRAXFA - PRINCIPAL AXIS FACTOR LOADINGS COMPUTATION
QTRANS - SQUARE ROOT FACTOR ANALYSIS OF A CORRELATION MATRIX
VARROT - VARIMAX ROTATION

3.3.3. DOCUMENTATION REFERENCES

WARD, J.H., JR., BUCHHORN, J., & HALL, K. INTRODUCTION TO PERSUB. PRL-TR-67-3(I), AD-660 578. LACKLAND AFB, TX: PERSONNEL RESEARCH LABORATORY, AEROSPACE MEDICAL DIVISION, AUGUST 1967.

WARD, J.H., JR., HALL, K., & BUCHHORN, J. PERSUB REFERENCE MANUAL. PRL-TR-67-3(II), AD-660 579. LACKLAND AFB, TX: PERSONNEL RESEARCH LABORATORY, AEROSPACE MEDICAL DIVISION, AUGUST 1967.

THE FIRST REFERENCE GIVEN ABOVE CONTAINS FOUR EXAMPLES OF THE APPLICATION OF THE PERSUB SYSTEM TO DATA ANALYSIS PROBLEMS. THE SECOND REFERENCE CONTAINS DESCRIPTIONS OF 49 PERSUB SUBROUTINES ALONG WITH A LISTING OF THE SOURCE LANGUAGE STATEMENTS FOR EACH SUBROUTINE. SUBROUTINES ADDED TO PERSUB AFTER PUBLICATION OF THE

MATHEMATICAL AND STATISTICAL SOFTWARE INDEX

ABOVE TECHNICAL REPORTS ARE DESCRIBED IN THE TWO AFWAL UNIVAC 1108 DOC PROCESSOR REPORTS GIVEN BELOW. THE FIRST CONTAINS DESCRIPTIONS OF MDRAND, MUTRUN, RANDES, RANERP, OBTAFT, PGRAPH, AND SECANT. THE SECOND REPORT DESCRIBES ARRAYT, COMMUN, COMPAR, CURVLD, MISC, PREDSC, AND TABDEV.

AFWAL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE = MISC.

AFWAL UNIVAC 1108 DOC PROCESSOR REPORT (UNPUBLISHED): TITLE = PERSUB.

MISC REPORT RETRIEVAL COMMAND: @T+T.DOC MISC.

PERSUB REPORT RETRIEVAL COMMAND: @T+T.DOC PERSUB.

Descriptive and Inferential Statistics

Descriptive Statistics

Descriptive statistics are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data. Descriptive statistics are typically distinguished from inferential statistics. With descriptive statistics you are simply describing what is or what the data shows. With inferential statistics, you are trying to reach conclusions that extend beyond the immediate data alone. For instance, we use inferential statistics to try to infer from the sample data what the population might think. Or, we use inferential statistics to make judgments of the probability that an observed difference between groups is a dependable one or one that might have happened by chance in this study. Thus, we use inferential statistics to make inferences from our data to more general conditions; we use descriptive statistics simply to describe what's going on in our data.

Descriptive statistics are very important because if we simply presented our raw data it would be hard to visualize what the data was showing, especially if there was a lot of it. Descriptive statistics therefore enable us to present the data in a more meaningful way, which allows simpler interpretation of the data. For example, if we had the results of 100 pieces of students' coursework, we may be interested in the overall performance of those students. We would also be interested in the distribution or spread of the marks. Descriptive statistics allow us to do this. How descriptive statistics are typically used is that there is a measure of central tendency and a measure of spread. Measures of central tendency include the mean, median and mode, while measures of spread include the standard deviation, variance and range. Together with graphical analysis, we usually need to use a measure of both central tendency and spread to describe the data.

THE BIOMEDICAL COMPUTER PROGRAMS HAVE STATISTICAL TITLES AND ARE GROUPED TOGETHER ACCORDING TO THE FOLLOWING GENERAL STATISTICAL CATEGORIES OR CLASSES:

CLASS D - DATA DESCRIPTION AND TABULATION; DATA IN GROUPS; PLOTS AND HISTOGRAMS
CLASS F - FREQUENCY TABLES
CLASS M - MULTIVARIATE ANALYSIS
CLASS R - REGRESSION ANALYSIS
CLASS S - SPECIAL PROGRAMS
CLASS T - TIME SERIES ANALYSIS
CLASS V - VARIANCE ANALYSIS.

THE FOLLOWING SECTION CONTAINS A LIST OF THE 30 BMDP PROGRAMS AVAILABLE IN THE SOFTWARE LIBRARY ALONG WITH THEIR DESCRIPTIONS FROM THE BMDP-77 MANUAL. IN EACH SECTION FOLLOWING THE BMDP PROGRAM DESCRIPTIONS ARE THE DESCRIPTIONS OF THE SOFTWARE LIBRARY'S AMD AND BMDX PROGRAMS. (NOTE. THE LETTER "P" FOLLOWS THE ACRONYM BMD IN THE PROGRAM NAME FOR "P" SERIES PROGRAMS, WHILE THE LETTER "X" IS USED FOR THE "X" SERIES PROGRAMS. ALSO, THE LAST LETTER IN THE PROGRAM NAMES BELOW CORRESPOND TO THE STATISTICAL CATEGORIES TO WHICH THE PROGRAMS BELONG.)

9.1.2. LIST OF PROGRAMS AND PROGRAM DESCRIPTIONS

CLASS D - DATA DESCRIPTION AND TABULATION; DATA IN GROUPS; PLOTS AND HISTOGRAMS

BMDP1D — SIMPLE DATA DESCRIPTION
COMPUTES UNIVARIATE STATISTICS, LISTS ALL DATA OR LISTS SELECTED CASES (THOSE CONTAINING MISSING VALUES OR VALUES OUTSIDE MINIMUM OR MAXIMUM LIMITS) AND STORES THE DATA IN A BMDP FILE. FOR EACH VARIABLE, THE OUTPUT INCLUDES MEAN, STANDARD DEVIATION, STANDARD ERROR OF MEAN, COEFFICIENT OF VARIATION, LARGEST AND SMALLEST VALUES, LARGEST AND SMALLEST STANDARDIZED SCORES (Z-SCORES), RANGE, AND THE TOTAL NUMBER OF ACCEPTABLE VALUES. CODES (VALUES) OR INTERVALS (CATEGORIES) FOR A VARIABLE CAN BE SPECIFIED AND FOR EACH SUCH VARIABLE P1D PRINTS THE NUMBER OF CASES WITH EACH CODE OR IN EACH INTERVAL.

BMDP2D — DETAILED DATA DESCRIPTION
COUNTS AND LISTS DISTINCT VALUES OF EACH VARIABLE IN ASCENDING ORDER, AND COMPUTES UNIVARIATE STATISTICS. VALUES OF VARIABLES MAY BE TRUNCATED OR ROUNDED BY DIFFERENT AMOUNTS FOR DIFFERENT VARIABLES. ALL COMPUTATIONS ARE PERFORMED ON THE TRUNCATED OR ROUNDED VALUES. FOR EACH VARIABLE, THE PROGRAM PRINTS THE NUMBER OF ACCEPTABLE VALUES, THE MAXIMUM, MINIMUM, RANGE, HALF OF THE INTERQUARTILE RANGE, MEAN, MEDIAN, MODE, STANDARD DEVIATION, STANDARD ERROR OF THE MEAN, SKEWNESS AND KURTOSIS. THE COMPUTED STATISTICS ARE PRINTED ON A LINE PLOT (130 CHARACTERS LONG); A HISTOGRAM IS ALSO PRINTED. FOR EACH DISTINCT VALUE, THE NUMBER OF CASES WITH EACH VALUE, THE PERCENT OF THE CASES COUNTED THAT THIS NUMBER REPRESENTS AND THE CUMULATIVE PERCENT OF CASES WITH VALUES LESS THAN OR EQUAL TO EACH DISTINCT VALUE ARE PRINTED. OPTIONAL OUTPUT INCLUDES THREE NEW LOCATION ESTIMATES.

BMDP3D — COMPARISON OF TWO GROUPS WITH T TESTS
COMPUTES ONE-SAMPLE AND TWO-SAMPLE T TESTS AND THEIR ASSOCIATED PROBABILITY LEVELS. CRUDE HISTOGRAMS, SHOWING CASE DISTRIBUTIONS, ARE PRESENTED WITH EACH TEST. A GROUPING VARIABLE THAT CLASSI-

FIES THE CASES INTO GROUPS CAN BE SPECIFIED. IF THERE ARE MORE THAN TWO GROUPS, THE PROGRAM COMPUTES TWO-SAMPLE T TESTS FOR EACH PAIR OF GROUPS; TESTS ARE COMPUTED FOR BOTH POOLED AND SEPARATE VARIANCE ESTIMATES. EQUALITY OF GROUP VARIANCES IS TESTED. IF SEVERAL VARIABLES ARE ANALYZED, MAHALANOBIS' D-SQUARE AND HOTELLING'S T-SQUARE CAN BE REQUESTED. A ONE-SAMPLE T TEST [H(0): MU = 0] IS COMPUTED WHEN THERE IS NO GROUPING VARIABLE. A ONE-SAMPLE T TEST [H(0): MU = MU(0)] OR A MATCHED PAIR T TEST MAY BE PERFORMED. THE OUTPUT INCLUDES MEANS, STANDARD DEVIATIONS, STANDARD ERROR OF MEANS, MAXIMUMS AND MINIMUMS FOR EACH GROUP. AN F VALUE FOR COMPARISON OF VARIANCES, POOLED AND SEPARATE T VALUES, AND TWO-TAILED PROBABILITY VALUES FOR EACH T AND F ARE ALSO PRINTED.

BMDP9D — SINGLE COLUMN FREQUENCIES
COUNTS THE NUMERIC AND NON-NUMERIC CHARACTERS (SYMBOLS) FOUND IN SINGLE COLUMN FIELDS. THE PROGRAM ASSUMES THAT A SINGLE COLUMN FIELD IS READ FOR EACH VARIABLE IN A1 FORMAT. ALL KEYPUNCH CHARACTERS ARE CONSIDERED LEGAL, AND THE FREQUENCY OF EACH IS COUNTED SEPARATELY. P9D CAN BE USED TO LIST THE DATA AFTER CERTAIN CHARACTERS ARE REPLACED BY SPECIFIED SYMBOLS (E.G., NUMBERS REPLACED BY BLANKS). THIS IS USEFUL FOR PRELIMINARY DATA SCREENING, VERIFYING APPROPRIATE CODING, AND VERIFYING (AT LEAST ROUGHLY) THE KINDS OF DATA PRESENT.

BMDP5D — HISTOGRAMS AND UNIVARIATE PLOTS
PRINTS HISTOGRAMS AND OTHER UNIVARIATE PLOTS. FOR EACH PLOT, CASES BELONGING TO ONE OR MORE GROUPS CAN BE USED. IN THE PLOTS, CASES FROM DIFFERENT GROUPS CAN BE IDENTIFIED BY DISTINCT LETTERS. PLOTS CAN BE PRINTED FOR ALL GROUPS IN ONE PLOT OR FOR EACH GROUP INDIVIDUALLY. THE SIZE OF THE PLOTS CAN BE SPECIFIED. ONE OR MORE OF THE FOLLOWING PLOTS ARE PRINTED FOR A VARIABLE: A HISTOGRAM OR A CUMULATIVE HISTOGRAM WITH FREQUENCIES AND PERCENTAGES IN EACH INTERVAL, A NORMAL PROBABILITY PLOT, A HALF-NORMAL PROBABILITY PLOT, OR A CUMULATIVE DISTRIBUTION PLOT. THE TOTAL FREQUENCY COUNT, MEAN AND STANDARD DEVIATION (FOR EACH GROUP IN A PLOT) ARE PRINTED.

BMDP6D — BIVARIATE (SCATTER) PLOTS
PRINTS BIVARIATE SCATTER PLOTS. CASES BELONGING TO ONE OR MORE GROUPS CAN BE USED FOR EACH PLOT. THE PLOTS MAY HAVE FREQUENCIES PLOTTED FOR EACH

POINT, OR DISTINCT LETTERS MAY BE USED TO INDICATE GROUP MEMBERSHIP. PLOTS CAN BE PRINTED FOR ALL GROUPS IN ONE PLOT OR FOR EACH GROUP IN SEPARATE PLOTS. PLOTS CAN BE SUPERIMPOSED ONE ON THE OTHER. THE SIZE OF THE PLOTS CAN BE CONTROLLED AND THE REGRESSION EQUATIONS FOR REGRESSING X ON Y AND Y ON X CAN BE REQUESTED; INTERSECTIONS OF THE REGRESSION LINES WITH THE PLOT FRAME ARE INDICATED. OPTIONAL STATISTICS PRINTED FOR EACH PLOT (FOR CASES USED IN THE PLOT) ARE MEAN AND STANDARD DEVIATION FOR EACH VARIABLE, AND EQUATION OF SIMPLE LINEAR REGRESSION FOR EACH VARIABLE ON THE OTHER VARIABLE WITH THEIR RESIDUAL MEAN SQUARE, CORRELATION, AND FREQUENCY OF CASES.

BMDP7D — DESCRIPTION OF GROUPS (STRATA) WITH HISTOGRAMS AND ANALYSIS OF VARIANCE
GROUPS THE DATA INTO A SPECIFIED NUMBER OF GROUPS BASED ON THE VALUE OF A GROUPING VARIABLE. FOR EACH VARIABLE, THE PROGRAM PRINTS HISTOGRAMS FOR EACH GROUP, SIDE BY SIDE. THE NUMBER OF INTERVALS IN THE HISTOGRAMS CAN BE SPECIFIED OR COMPUTED BY THE PROGRAM. MEAN, STANDARD DEVIATION, AND FREQUENCIES ARE REPORTED FOR EACH GROUP AND FOR ALL GROUPS COMBINED. A ONE-WAY OR TWO-WAY ANALYSIS OF VARIANCE IS COMPUTED. SELECTED VALUES FOR EACH VARIABLE CAN BE EXCLUDED FROM THE COMPUTATIONS BY SPECIFYING MISSING VALUE CODES, MAXIMUMS OR MINIMUMS; HOWEVER, THESE VALUES ARE DISPLAYED SEPARATELY IN THE HISTOGRAM. OPTIONAL OUTPUT INCLUDES A LISTING OF INPUT DATA AFTER TRANSFORMATIONS OR AFTER ORDERING FROM LOW TO HIGH ON A SPECIFIED VARIABLE; CORRELATIONS FOR EACH GROUP SEPARATELY AND FOR ALL GROUPS COMBINED; AND WINSORIZED MEANS AND 95% CONFIDENCE INTERVALS FOR THE WINSORIZED MEANS FOR EACH GROUP.

BMDP8D — MISSING VALUE CORRELATION
COMPUTES CORRELATIONS FOUR DIFFERENT WAYS FROM DATA CONTAINING MISSING VALUES. THESE COMPUTATIONS USE DIFFERENT CASES IN COMPUTING MEANS AND VARIANCES. COMPUTATIONS CAN BE PERFORMED USING ALL ACCEPTABLE VALUES -- THE MEANS ARE COMPUTED FROM ALL ACCEPTABLE VALUES AND THEN DEVIATIONS FROM THESE MEANS ARE USED TO COMPUTE COVARIANCES AND CORRELATIONS; ALL ACCEPTABLE PAIRS OF VALUES FOR COVARIANCES -- EACH ELEMENT OF THE COVARIANCE MATRIX IS COMPUTED FROM THE EXISTING PAIRS OF VALUES INVOLVED, WHICH ARE THEN USED TO COMPUTE THE CORRELATIONS; ALL ACCEPTABLE PAIRS OF VALUES

FOR CORRELATIONS -- EACH ELEMENT OF THE CORRELATION MATRIX IS COMPUTED FROM THE ACCEPTABLE PAIRS OF VALUES INVOLVED; OR ALL COMPLETE CASES -- CASES WITH ANY EXCLUDED CASES ARE NOT USED. ANY OR ALL OF THE FIRST THREE TYPES OR THE FOURTH TYPE CAN BE REQUESTED FOR EACH PROBLEM. CASE WEIGHTS MAY ALSO BE SPECIFIED. THE CORRELATION MATRIX CAN BE SAVED IN A BMDP FILE AND USED AS INPUT TO OTHER PROGRAMS. P8D CAN ALSO BE USED TO OBTAIN A SUBMATRIX OF THE CORRELATION MATRIX. THE OUTPUT CAN INCLUDE THE MEAN AND VARIANCE OF EACH VARIABLE, A FREQUENCY TABLE OF ACCEPTABLE VALUES FOR PAIRS OF VARIABLES, THE COVARIANCE MATRIX, THE CORRELATION MATRIX, THE SUM OF WEIGHTS, THE MATRIX OF MEANS AND THE MATRIX OF VARIANCES.

BMDP9D — MULTIWAY DESCRIPTION OF GROUPS

PROVIDES CELLWISE DESCRIPTIVE STATISTICS FOR CASES WHEN THE DATA ARE CLASSIFIED SIMULTANEOUSLY BY SEVERAL GROUPING VARIABLES. THE PROGRAM GIVES CELL FREQUENCIES, MEANS AND STANDARD DEVIATIONS FOR EACH VARIABLE DESCRIBED. IF REQUESTED, THE PROGRAM YIELDS THE SAME INFORMATION FOR SPECIFIED MARGINAL CELLS. P9D COMPUTES A CHI-SQUARE TEST FOR EQUALITY OF ALL CELL FREQUENCIES, AN F TEST FOR EQUALITY OF ALL CELL MEANS (USING ONLY NONEMPTY CELLS) AND BARTLETT'S TEST FOR HOMOGENEITY OF VARIANCES (USING ALL CELLS WITH NONZERO VARIANCE). FOR EACH VARIABLE DESCRIBED, THE PROGRAM ALSO PRINTS A PROFILE PLOT INDICATING HOW CELL MEANS SHIFT FROM CELL TO CELL. IN THESE PLOTS THE NUMBERS PRINTED IN THE PLOT INDICATE CELL FREQUENCIES; THE LOCATION OF THE NUMBERS INDICATES THE DEVIATION OF CELL MEANS FROM THE GRAND MEAN IN STANDARD DEVIATION UNITS.

BMDX70 (OR BMD13D) — T PROGRAM

COMPUTES T-STATISTICS AND ASSOCIATED PROBABILITY LEVELS FOR THE EQUALITY OF THE MEANS OF TWO GROUPS BASED ON POOLED AND SEPARATE VARIANCE ESTIMATES. AN F-STATISTIC AND ASSOCIATED PROBABILITY LEVEL FOR THE EQUALITY OF GROUP VARIANCES ARE ALSO COMPUTED.

BMDX84 (OR BMD12D) — ASYMMETRIC CORRELATION WITH MISSING DATA

COMPUTES LARGE CORRELATION MATRICES OR PORTIONS OF LARGE CORRELATION MATRICES FROM DATA WITH MISSING VALUES, USING ALL POSSIBLE COMPLETE PAIRS WITH BLANKS DENOTING MISSING DATA.

CLASS F - FREQUENCY TABLES

BMDP1F TWO-WAY FREQUENCY TABLES

PRINTS A VARIETY OF TWO-WAY TABLES AND PROVIDES ESTIMATES OF MANY MEASURES OF ASSOCIATION AND PREDICTION. THE TABLES CAN CONTAIN OBSERVED FREQUENCIES; EXPECTED VALUES UNDER INDEPENDENCE; PERCENTAGES OF THE TABLE TOTAL, OF EACH ROW TOTAL OR OF EACH COLUMN TOTAL; OR SEVERAL TYPES OF RESIDUALS. TESTS AND STATISTICS THAT CAN BE COMPUTED FOR EACH TWO-WAY TABLE ARE: TESTS OF INDEPENDENCE AND RELATED MEASURES (CHI-SQUARE TEST, THE LIKELIHOOD RATIO CHI-SQUARE TEST G-SQUARE, PHI, THE CONTINGENCY COEFFICIENT AND CRAMER'S V); TESTS OF INDEPENDENCE APPROPRIATE ONLY FOR 2X2 TABLES (FISHER'S EXACT TEST AND YATE'S CORRECTED CHI-SQUARE); MEASURES OF ASSOCIATION APPROPRIATE FOR A 2X2 TABLE (YULE'S Q AND Y, THE CROSS-PRODUCT RATIO AND THE TETRACHORIC CORRELATION R(T)); MEASURES OF ASSOCIATION AND CORRELATION WHEN THE CATEGORIES OF BOTH FACTORS ARE ORDERED (GOODMAN AND KRUSKAL GAMMA, KENDALL'S TAU(B), STUART'S TAU(C), THE PRODUCT-MOMENT CORRELATION R AND THE SPEARMAN RANK CORRELATION); PREDICTIVE MEASURES WHEN THE CATEGORIES OF BOTH FACTORS ARE ORDERED (SOMERS' D); PREDICTIVE MEASURES WHEN EITHER OR BOTH FACTORS ARE QUALITATIVE (THE GOODMAN AND KRUSKAL TAU, LAMBDA AND LAMBDA(•) AND THE UNCERTAINTY COEFFICIENT); MCNEMAR'S TEST OF SYMMETRY; AND A TEST OF A CONTRAST ON THE PROPORTIONS IN A 2XK TABLE. STANDARD ERRORS ARE PROVIDED FOR MANY OF THE MEASURES OF ASSOCIATION AND PREDICTION.

BMDP2F TWO-WAY FREQUENCY TABLES -- EMPTY CELLS AND DEPARTURES FROM INDEPENDENCE

TREATS TWO SPECIAL PROBLEMS: FIRST, CELLS CAN BE EXCLUDED FROM THE FREQUENCY TABLE, SUCH AS CELLS WITH A PRIOR PROBABILITY OF ZERO (STRUCTURAL ZEROES). P2F COMPUTES THE APPROPRIATE TEST OF INDEPENDENCE BETWEEN ROWS AND COLUMNS (THE CHI-SQUARE TEST OF QUASI-INDEPENDENCE). THIS TEST USES ALL CELLS THAT ARE NOT EXCLUDED A PRIORI AND IS IDENTICAL TO THE USUAL CHI-SQUARE TEST WHEN THERE ARE NO EXCLUDED CELLS. SECOND, CELLS OR PATTERNS OF CELLS CAN BE IDENTIFIED THAT CONTRIBUTE TO A DEPARTURE FROM INDEPENDENCE OR QUASI-INDEPENDENCE; I.E., CONTRIBUTE GREATLY TO A SIGNIFI-

CANT CHI-SQUARE TEST. A CRITERION CAN BE SELECTED BY WHICH CELLS ARE IDENTIFIED AND ELIMINATED (TREATED AS EXCLUDED) FROM THE TABLE IN A STEPWISE MANNER. ONE SUCH CRITERION IS TO ELIMINATE THE CELL THAT, WHEN EXCLUDED, MINIMIZES THE CHI-SQUARE (OR LIKELIHOOD RATIO CHI-SQUARE) TEST OF QUASI-INDEPENDENCE PERFORMED ON ALL CELLS NOT PREVIOUSLY EXCLUDED. INFERENCES CAN BE DRAWN FROM THE DATA BY EXAMINING THE PATTERN OF CELLS ELIMINATED BY THIS PROCEDURE. P2F COMPUTES ONLY THE CHI-SQUARE TESTS OF INDEPENDENCE (OR QUASI-INDEPENDENCE), EXPECTED VALUES ASSUMING INDEPENDENCE, AND VARIOUS TYPES OF RESIDUALS. IT DOES NOT PROVIDE ESTIMATES OF THE MEASURES OF ASSOCIATION AND PREDICTION THAT ARE AVAILABLE IN P1F.

BMDP3F — MULTIWAY FREQUENCY TABLES

ANALYZES DATA IN A MULTIWAY TABLE. THE PURPOSE OF THE ANALYSIS IS TO OBTAIN A DESCRIPTION OF THE RELATIONSHIPS BETWEEN THE FACTORS OF THE TABLE, EITHER BY FORMING A MODEL FOR THE DATA OR BY TESTING AND ORDERING THE IMPORTANCE OF THE INTERACTIONS BETWEEN THE FACTORS. THE ANALYSIS IS BASED ON FITTING A (HIERARCHICAL) LOG-LINEAR MODEL TO THE CELL FREQUENCIES; THAT IS, THE LOGARITHM OF THE EXPECTED CELL FREQUENCY IS WRITTEN AS AN ADDITIVE FUNCTION OF MAIN EFFECTS AND INTERACTIONS IN A MANNER SIMILAR TO THE USUAL ANALYSIS OF VARIANCE MODEL. IT SHOULD BE NOTED THAT THE VARIOUS TESTS FOR THE INTERACTIONS ARE NOT INDEPENDENT. IF A TWO-WAY TABLE WERE INPUT TO P3F, ONE POSSIBLE LOG-LINEAR MODEL WOULD INCLUDE TWO MAIN EFFECTS AND A SINGLE INTERACTION. THE TEST THAT THE INTERACTION IS ZERO IS IDENTICAL TO THE TEST FOR INDEPENDENCE OBTAINED IN P1F. P3F CAN BE USED TO FIND AN APPROPRIATE MODEL FOR THE DATA IN THE TABLE BY FITTING SPECIFIED MODELS OR BY USING A METHOD TO SCREEN EFFECTS. THE PROGRAM TESTS THE APPROPRIATENESS OF MODELS BY THE LIKELIHOOD RATIO CHI-SQUARE (G-SQUARE) AND BY THE USUAL CHI-SQUARE GOODNESS-OF-FIT. P3F PRINTS THE TABLE OF OBSERVED FREQUENCIES AND CAN PRINT MARGINAL TABLES OF THE OBSERVED FREQUENCIES AS WELL. IT CAN TEST CERTAIN CLASSES OF MODELS OR SPECIFIED MODELS. FOR EACH SPECIFIED MODEL IT CAN PRINT THE EXPECTED VALUES AND TWO TYPES OF RESIDUALS, ESTIMATES OF THE PARAMETERS OF THE MODEL AND THEIR STANDARD ERRORS, AND TESTS-OF-FIT OF MODELS THAT DIFFER FROM THE SPECIFIED MODEL BY ONE EFFECT.

CLASS M - MULTIVARIATE ANALYSIS

BMDP1M CLUSTER ANALYSIS OF VARIABLES
CLUSTERS SIMILAR VARIABLES BY USING AN INITIAL MEASURE OF ASSOCIATION BETWEEN PAIRS OF VARIABLES (E.G., CORRELATION) TO FORM A CLUSTER OF THE TWO MOST SIMILAR VARIABLES AND THEN USING AN AMALGAMATION (LINKAGE) RULE TO FORM FURTHER CLUSTERS. THE AMALGATION RULE DETERMINES THE DEGREE OF ASSOCIATION BETWEEN ANY TWO CLUSTERS WHERE A CLUSTER CONSISTS OF ONE OR MORE VARIABLES. WHEN THE CLUSTERING PROCESS IS FINISHED, EACH CLUSTER CONSISTS OF TWO OR MORE VARIABLES AND EACH VARIABLE IS PLACED IN ONE OR MORE CLUSTERS. BOTH THE MEASURE OF ASSOCIATION FOR THE VARIABLES AND THE AMALGAMATION RULE FOR THE CLUSTERS MAY BE SPECIFIED. P1M ACCEPTS A DATA MATRIX OR A DISTANCE MATRIX (MEASURE OF ASSOCIATION) AS INPUT. OUTPUT INCLUDES A SUMMARY TABLE FOR THE CLUSTERING PROCESS, A TREE DIAGRAM OF CLUSTERS SUPERIMPOSED OVER A SIMILARITY OR DISTANCE MATRIX SCALED FROM 0 TO 100, A TABLE OF SIMILARITY OR DISTANCE MATRIX SCALING, AND AN EXPLANATION OF THE TREE DIAGRAM FOR THE FIRST PROBLEM. THE CORRELATION MATRIX CAN ALSO BE REPRESENTED IN SHADED FORM AFTER REARRANGING THE ORDER OF THE VARIABLES ACCORDING TO THE CLUSTERS.

BMDP2M CLUSTER ANALYSIS OF CASES
CLUSTERS CASES ACCORDING TO ONE OF FOUR AVAILABLE DISTANCE MEASURES. THE TWO CASES HAVING THE SHORTEST DISTANCE BETWEEN THEM ARE AMALGAMATED AND TREATED AS ONE CASE AND THEN, IN TURN, CLUSTERED WITH OTHERS. THIS ALGORITHM CONTINUES UNTIL ALL CASES AND CLUSTERS ARE AMALGAMATED INTO ONE CLUSTER. THE DISTANCE BETWEEN CASES IS EITHER THE P-TH ROOT OF THE SUMS OF THE P-TH POWERS OF DIFFERENCES (WHEN P=2 THIS IS THE EUCLIDEAN DISTANCE) OR CHI-SQUARE OR PHI-SQUARE (=CHI-SQUARE/N) WHERE CHI-SQUARE IS COMPUTED FROM THE TABLE FORMED BY THE TWO CASES AND THE VARIABLES (THE LAST TWO DISTANCE MEASURES ARE USEFUL WHEN THE DATA ARE COUNTS). A DIAGRAM DRAWN WITH VERTICAL LINES TO INDICATE CLUSTERING OF THE CASES IS PRINTED, AND THE ORDER OF CLUSTERING IS INDICATED IN THE DIAGRAM. OPTIONAL OUTPUT INCLUDES THE INPUT DATA MATRIX AFTER STANDARDIZATION, THE INITIAL DISTANCE MATRIX BETWEEN CASES, AND A HORIZONTAL CLUSTERING TREE DIA-

GRAM. THE DISTANCE MATRIX CAN BE PRINTED IN SHADED FORM AFTER REARRANGING THE CASES ACCORDING TO THE CLUSTERS.

BMDP3M — BLOCK CLUSTERING
SIMULTANEOUSLY CLUSTERS CASES AND VARIABLES OF A DATA MATRIX. BLOCKS (SUBMATRICES OF THE DATA MATRIX) ARE IDENTIFIED WHILE THE MARGINAL ROWS OF A BLOCK FORM A CLUSTER OF CASES, AND THE MARGINAL COLUMNS OF A BLOCK FORM A CLUSTER OF VARIABLES. THREE TYPES OF CLUSTERS ARE THUS DISCOVERED: CLUSTERS OF DATA VALUES (THE BLOCKS), CLUSTERS OF CASES, AND CLUSTERS OF VARIABLES. A BLOCK DIAGRAM COMPLETELY REPRESENTS THE DATA, WITH RELATIVELY FEW PRINTED SYMBOLS. IN THE BLOCK, THE ROWS (CASES) AND COLUMNS (VARIABLES) OF DATA HAVE BEEN PERMUTED, AND SMALLER BLOCKS (OR SUBMATRICES) OF SIMILAR VALUES ARE OUTLINED. SIMILAR VALUES ARE LEFT BLANK TO GIVE A GOOD VISUAL REPRESENTATION OF LIKE VALUES. TO PREVENT OVERLAPPING, THE BLOCKS FORM A HIERARCHY (THAT IS, ANY PAIR OF BLOCKS ARE DISJOINT, OR ONE INCLUDES THE OTHER). AT AN EARLY STAGE OF DATA ANALYSIS, THE PROGRAM CAN BE USED AS A MULTIVARIATE HISTOGRAM. IN ONE-DIMENSIONAL HISTOGRAMS REAL-VALUED VARIABLES ARE SEPARATED INTO CATEGORIES OF EQUAL LENGTH. IN A SENSE, P3M DOES THIS TOO. IT SEPARATES THE VALUES OF EACH VARIABLE INTO A MAXIMUM OF 35 INTERVALS AND CODES THEM 1,2,...,9,A,B,...,Z. FOR EACH VARIABLE THE NUMBER OF CATEGORIES, OR INTERVALS USED, CAN BE SPECIFIED. THE OUTPUT REPORTS THE NUMBER OF TIMES EACH VARIABLE TAKES EACH OF ITS CODED VALUES. TWO TREES ARE PRINTED, ONE REPORTING THE JOINING SEQUENCE FOR CASES (ROW CLUSTERS) AND THE OTHER REPORTING THE JOINING SEQUENCE FOR VARIABLES (COLUMN CLUSTERS). ALSO INCLUDED IS A BLOCK DIAGRAM ON THE PERMUTED DATA MATRIX, IN WHICH A LEADER VALUE AND EXCEPTIONS ARE PRINTED FOR EACH BLOCK. VALUES SIMILAR TO THE LEADER ARE BLANKED OUT.

BMDP4M — FACTOR ANALYSIS
PERFORMS A FACTOR ANALYSIS OF EITHER A CORRELATION OR A COVARIANCE MATRIX. THE INPUT CAN BE DATA, CORRELATION MATRIX, COVARIANCE MATRIX, FACTOR LOADINGS OR FACTOR SCORE COEFFICIENTS. SEVERAL INITIAL COMMUNALITY ESTIMATES CAN BE USED, INCLUDING SQUARED MULTIPLE CORRELATIONS, MAXIMUM ROW VALUES AND VALUES SPECIFIED BY THE USER. INITIAL FACTOR EXTRACTION CAN BE OBTAINED BY PRINCIPAL COMPONENTS, MAXIMUM LIKELIHOOD, KAISER'S SECOND GENERATION LITTLE JIFFY, OR

ITERATED PRINCIPAL FACTOR ANALYSIS. SEVERAL METHODS OF ROTATION ARE AVAILABLE, INCLUDING VARIMAX AND DIRECT QUARTIMIN. FACTOR SCORES ARE COMPUTED FOR EACH CASE, AND MAHALANOBIS DISTANCES ARE COMPUTED FROM THE CENTROID OF ALL CASES FOR THE FACTOR SCORES, RAW DATA, AND THE RESIDUALS OF THE ORIGINAL DATA REGRESSED ON THE FACTOR SCORES. OUTPUT INCLUDES MEANS, STANDARD DEVIATIONS, COEFFICIENTS OF VARIATION, MAXIMUMS AND MINIMUMS, STANDARD SCORES, CORRELATION MATRIX, COVARIANCE MATRIX, FACTOR LOADINGS AND THEIR PLOTS, SQUARED MULTIPLE CORRELATIONS, FACTOR CORRELATIONS, FACTOR STRUCTURE, FACTOR SCORES, FACTOR SCORE COEFFICIENTS, FACTOR SCORE COVARIANCES AND FACTOR SCORE PLOTS. THE CORRELATION MATRIX CAN BE PRINTED IN SHADED FORM AFTER REARRANGING THE ORDER OF THE VARIABLES ACCORDING TO THE FACTOR LOADINGS.

BMDP6M — CANONICAL CORRELATION ANALYSIS

COMPUTES CANONICAL CORRELATION ANALYSIS FOR TWO SETS OF VARIABLES. INPUT CAN BE A DATA MATRIX, COVARIANCE MATRIX OR A CORRELATION MATRIX. OUTPUT INCLUDES MEANS, STANDARD DEVIATIONS, COEFFICIENTS OF VARIATION, SKEWNESS, KURTOSIS, MINIMUMS, MAXIMUMS, CORRELATIONS, COVARIANCES, CANONICAL CORRELATIONS, EIGENVALUES ASSOCIATED WITH EACH PAIR OF CANONICAL VARIABLES AND BARTLETT'S TEST FOR THE SIGNIFICANCE OF THE REMAINING EIGENVALUES, COEFFICIENTS FOR CANONICAL VARIABLES, THE SCORES OF THE CANONICAL VARIABLES FOR EACH CASE, CORRELATIONS OF VARIABLES WITH CANONICAL VARIABLES (LOADINGS) AND BIVARIATE PLOTS FOR ORIGINAL VARIABLES AND CANONICAL VARIABLES.

BMDP7M — STEPWISE DISCRIMINANT ANALYSIS

PERFORMS A MULTIPLE GROUP DISCRIMINANT ANALYSIS. THE VARIABLES USED IN COMPUTING THE LINEAR CLASSIFICATION FUNCTIONS ARE CHOSEN IN A STEPWISE MANNER. AT EACH STEP THE VARIABLE THAT ADDS MOST TO THE SEPARATION OF THE GROUPS IS ENTERED. BY SPECIFYING CONTRASTS GROUP DIFFERENCES OF INTEREST MAY BE STATED. THESE CONTRASTS GUIDE SELECTION OF THE VARIABLES. FOR EACH CASE THE GROUP CLASSIFICATIONS ARE EVALUATED. BASED ON THE POSTERIOR PROBABILITIES, A CLASSIFICATION TABLE IS COMPUTED (PRIOR PROBABILITIES CAN BE SPECIFIED FOR USE IN THESE COMPUTATIONS). IN ADDITION A JACKKNIFE-VALIDATION PROCEDURE CAN BE REQUESTED TO REDUCE THE BIAS IN THE GROUP CLASSIFICATIONS. THE PROGRAM COMPUTES CANONICAL DISCRIMINANT FUNCTIONS AND PLOTS THE FIRST TWO TO GIVE AN OPTIMAL TWO-DIMENSIONAL PICTURE OF THE SEPARATION OF THE

GROUPS. OUTPUT INCLUDES MEANS, STANDARD DEVIATIONS, F STATISTICS, AND DEGREES OF FREEDOM FOR EACH VARIABLE AT EACH STEP, F STATISTICS FOR DISTANCES BETWEEN PAIRS OF GROUPS, WILKS' LAMBDA (U STATISTIC) FOR MULTIVARIATE ANOVA, MAHALANOBIS' D-SQUARE OF EACH CASE FROM EACH GROUP MEAN, CLASSIFICATION FUNCTIONS AND MATRICES, AND PERCENT CORRECT CLASSIFICATION, COEFFICIENTS FOR CANONICAL DISCRIMINANT FUNCTIONS, CANONICAL CORRELATIONS, CANONICAL VARIABLES AND A PLOT OF THE FIRST TWO CANONICAL VARIABLES.

BMDP4M — DESCRIPTION AND ESTIMATION OF MISSING VALUES
DESCRIBES THE PATTERN OF MISSING VALUES FOR MULTIVARIATE DATA AND PROVIDES ESTIMATES OF THE MISSING DATA. ESTIMATES CAN BE MEANS; PREDICTED VALUES FOR EACH MISSING VARIABLE DERIVED FROM REGRESSING IT ON THE AVAILABLE VARIABLE WITH WHICH IT IS MOST HIGHLY HIGHLY CORRELATED; PREDICTED VALUES FOR EACH MISSING VARIABLE DERIVED FROM ITS STEPWISE REGRESSION ON SOME OF THE AVAILABLE VARIABLES; AND PREDICTED VALUES FOR EACH MISSING VARIABLE DERIVED FROM REGRESSING IT ON ALL AVAILABLE VARIABLES. THE COVARIANCE MATRIX CAN BE COMPUTED USING COMPLETE CASES ONLY, OR USING ALL ACCEPTABLE VALUES. WHEN THE LATTER METHOD IS USED OR EXPLICITLY REQUESTED, THE EIGENVALUES ARE FOUND, AND THE CORRELATION MATRIX IS REESTIMATED USING ONLY POSITIVE EIGENVALUES AND THEIR EIGENVECTORS. A GROUPING VARIABLE AND CASE WEIGHTS CAN BE SPECIFIED. THE COVARIANCE MATRIX AND THE DATA WITH ESTIMATES REPLACING THE MISSING VALUES CAN BE SAVED IN A BMDP FILE. THE MATRIX OF PAIRWISE FREQUENCIES OF VARIABLES CAN ALSO BE SAVED AND ANALYZED IN OTHER BMDP PROGRAMS (SUCH AS FACTOR ANALYSIS PROGRAM, BMDP4M) TO FURTHER STUDY THE PATTERN OF MISSING VALUES. P4M CAN PRINT THE DATA, THE PATTERN OF MISSING DATA, SQUARED MULTIPLE CORRELATIONS OF EACH VARIABLE WITH ALL OTHER VARIABLES, MISSING DATA ESTIMATES AND MAHALANOBIS DISTANCES FOR EACH CASE TO THE CENTROID OF ALL CASES, BIVARIATE PLOTS FOR SPECIFIED PAIRS OF VARIABLES, THE CORRELATION MATRIX, EIGENVALUES OF THE CORRELATION MATRIX, AND THE COVARIANCE MATRIX. BIVARIATE SCATTER PLOTS CAN BE REQUESTED FOR ANY PAIR OF VARIABLES. IN THESE PLOTS ESTIMATED VALUES ARE DISTINGUISHED FROM THE ORIGINAL VALUES. FOR EACH VARIABLE IN EACH GROUP P4M PRINTS THE SAMPLE SIZE, PERCENTAGE MISSING, MEAN, STANDARD DEVIATION, COEFFICIENT OF VARIATION, MAXIMUM, MINIMUM, MAXIMUM STANDARD SCORE, MINIMUM STANDARD SCORE, SKEWNESS, AND KURTOSIS.

BMDO1M — PRINCIPAL COMPONENT ANALYSIS
COMPUTES PRINCIPAL COMPONENTS OF STANDARDIZED DATA, AND RANK ORDERS EACH STANDARDIZED CASE BY THE SIZE OF EACH PRINCIPAL COMPONENT SEPARATELY. OUTPUT INCLUDES CORRELATION COEFFICIENTS, EIGENVALUES AND RANK ORDER OF EACH STANDARDIZED CASE. WILL HANDLE UP TO 25 ORIGINAL VARIABLES AND 400 CASES.

BMD02M — REGRESSION ON PRINCIPAL COMPONENTS
COMPUTES PRINCIPAL COMPONENTS OF STANDARDIZED DATA AND RANK ORDERS EACH STANDARDIZED CASE BY THE SIZE OF EACH PRINCIPAL COMPONENT SEPARATELY. EACH DEPENDENT VARIABLE IS REGRESSED ON THE FIRST, FIRST TWO, FIRST THREE, AND ALL PRINCIPAL COMPONENTS WHEN EACH COMPONENT IS EXPRESSED IN TERMS OF STANDARDIZED DATA.

BMD03M — GENERAL FACTOR ANALYSIS
PERFORMS A PRINCIPAL COMPONENT SOLUTION AND ORTHOGONAL ROTATION OF THE FACTOR MATRIX. DATA INPUT MAY BE RAW DATA, A CORRELATION MATRIX, OR A FACTOR MATRIX. COMMUNALITIES ARE ESTIMATED FROM THE SQUARED MULTIPLE CORRELATION COEFFICIENTS OR THE MAXIMUM ABSOLUTE ROW VALUES, OR THEY MAY BE SPECIFIED BY THE USER TO REPLACE THE DIAGONAL OF THE CORRELATION MATRIX. WILL HANDLE 80 VARIABLES MEASURED ON 9,999 SUBJECTS.

BMD04M — DISCRIMINANT ANALYSIS FOR TWO GROUPS
COMPUTES A LINEAR FUNCTION OF P VARIABLES MEASURED ON EACH INDIVIDUAL OF TWO GROUPS; THIS CAN SERVE AS AN INDEX FOR DISCRIMINATION BETWEEN THE GROUPS. IT IS DETERMINED FROM THE CRITERION OF "BEST" IN THAT THE DIFFERENCE BETWEEN THE MEAN INDICES FOR THE TWO GROUPS DIVIDED BY A POOLED STANDARD DEVIATION OF THE INDICES IS MAXIMIZED (ESSENTIALLY HOTELLING'S T-SQUARE TEST ON TWO VECTORS OF MEANS). SEE ALSO BMD07M.

BMD05M — DISCRIMINANT ANALYSIS FOR SEVERAL GROUPS
DUPLICATES THE PROCEDURES OF BMD04M, BUT FOR TWO TO FIVE GROUPS WITH UP TO 25 VARIABLES AND NO GROUP EXCEEDING 175 IN NUMBER. INPUT DATA ARE A SET OF OBSERVATIONS FOR EACH OF THE CLASSIFICATION GROUPS AND EACH OBSERVATION CONTAINS A VALUE FOR EACH OF THE VARIABLES. THE COMPUTATIONAL PROCEDURE EVALUATES THE COMPUTED LINEAR FUNCTION CORRESPONDING TO EACH OF THE GROUPS AND ASSIGNS AN INDIVIDUAL TO THE GROUP FOR WHICH THE VALUE IS LARGEST. SEE ALSO

BMDIOM.

BMD10M CANONICAL CORRELATION ANALYSIS
COMPUTES THE CANONICAL CORRELATIONS BETWEEN TWO SETS OF VARIABLES. OUTPUT INCLUDES THE SIMPLE CORRELATION MATRIX, STANDARD DEVIATIONS, AND CANONICAL COEFFICIENTS FOR STANDARDIZED VARIABLES. WILL HANDLE UP TO A TOTAL OF 100 VARIABLES WITH THE SMALLER OF THE TWO SETS OF VARIABLES CONTAINING NO MORE THAN 20 VARIABLES.

BMD07M STEPWISE DISCRIMINANT ANALYSIS
PERFORMS A MULTIPLE DISCRIMINANT ANALYSIS IN A STEPWISE MANNER FOR 2 TO 20 GROUPS WITH UP TO 80 VARIABLES. A SET OF LINEAR CLASSIFICATION FUNCTIONS IS COMPUTED BY CHOOSING THE INDEPENDENT VARIABLES IN A STEPWISE MANNER, USING THESE FUNCTIONS AND PRIOR PROBABILITIES, THE POSTERIOR PROBABILITIES OF EACH CASE BELONGING TO EACH GROUP ARE COMPUTED. ALSO COMPUTES COEFFICIENTS FOR CANONICAL VARIABLES TO GIVE A TWO-DIMENSIONAL PLOT OF GROUP SEPARATIONS. GROUPS ARE EVALUATED IN PAIRS FOR SIGNIFICANCE OF DIFFERENCES OF MEANS BY THE F-RATIO AND MAHALANOBIS D-SQUARE.

BMD12D FACTOR ANALYSIS
(FOR BMDX85) PERFORMS FACTOR ANALYSIS ON A MATRIX OF LESS INPUT VARIABLES, USING EITHER COVARIANCE OR CORRELATION MATRICES. INITIAL COMMUNALITY ESTIMATES MAY BE SQUARED MULTIPLE CORRELATIONS, REPRESSION APPROACHES, MAXIMUM ABSOLUTE RIM VALUES, OR THEY MAY BE SPECIFIED BY THE USER. IF REQUESTED, THE PROGRAM WILL ITERATE ON THE INITIAL COMMUNALITY ESTIMATES. THREE TYPES OF ROTATION ARE AVAILABLE, ALL BASED ON THE VARIMAX CRITERIA. FACTOR SCORES MAY BE ESTIMATED. DATA INPUT MAY BE IN THE FORM OF RAW DATA, A CORRELATION MATRIX, A COVARIANCE MATRIX, OR A FACTOR LOADING MATRIX. THIS PROGRAM SUPERSEDES BMD12M.

BMD15D IDENTIFICATION OF OUTLIERS
(FOR BMDX85) DETERMINES MULTIVARIATE DATA FOR OUTLIERS BY COMPUTING THE MAHALANOBIS DISTANCE OF EACH CASE FROM THE CENTER OF THE DISTRIBUTION OF THE REMAINING CASES. IF THE PROBABILITY OF THE F-STATISTIC CORRESPONDING TO THE GREATEST DISTANCE IS SMALLER THAN A SPECIFIED VALUE, THE CASE INVOLVED IS OMITTED.

제5장 [illegible]

[illegible]

제1절 [illegible]

[illegible]

제2절 [illegible]

[illegible]

BMDP3R — NONLINEAR REGRESSION

OBTAINS A LEAST SQUARES FIT TO A NONLINEAR FUNCTION. IT IS APPROPRIATE FOR A WIDE VARIETY OF PROBLEMS THAT ARE NOT WELL-REPRESENTED BY EQUATIONS THAT ARE LINEAR IN THE PARAMETERS. FIVE OF THE MOST FREQUENTLY USED NONLINEAR FUNCTIONS ARE AVAILABLE BY SIMPLY STATING THEIR CODE NUMBER. OTHER FUNCTIONS CAN BE USED BY STATING THE FUNCTION AND ITS DERIVATIVES IN FORTRAN STATEMENTS. THE FIT IS OBTAINED BY MEANS OF GAUSS-NEWTON ITERATIONS. LINEAR EQUALITY CONSTRAINTS ON THE PARAMETERS CAN BE STATED, AND CASE WEIGHTS CAN BE USED. PARAMETERS CAN ALSO BE ESTIMATED BY MAXIMUM LIKELIHOOD. AFTER EACH ITERATION, THE PARAMETER VALUES AND RESIDUAL SUM OF SQUARES ARE PRINTED. AFTER THE LAST ITERATION, THE PROGRAM REPORTS THE ASYMPTOTIC CORRELATIONS AND STANDARD DEVIATIONS FOR THE ESTIMATED PARAMETERS. FOR EACH CASE AFTER THE LAST ITERATION, P3R LISTS THE PREDICTED AND OBSERVED VALUES FOR THE DEPENDENT VARIABLE, RESIDUAL VALUE, AND OBSERVED VALUES FOR THE INDEPENDENT VARIABLES. SCATTER PLOTS OF PREDICTED AND OBSERVED VALUES VERSUS SELECTED VARIABLE VALUES MAY BE REQUESTED, AS WELL AS RESIDUALS VERSUS SELECTED VARIABLES, AND NORMAL AND DETRENDED NORMAL PROBABILITY PLOTS OF RESIDUALS.

BMDP4R — REGRESSION ON PRINCIPAL COMPONENTS

COMPUTES A REGRESSION ANALYSIS FOR EACH DEPENDENT VARIABLE ON A SET OF PRINCIPAL COMPONENTS COMPUTED FROM THE INDEPENDENT VARIABLES. THE PRINCIPAL COMPONENTS ARE COMPUTED FROM THE ORIGINAL VARIABLES (USING THE COVARIANCE MATRIX) OR THE STANDARDIZED VARIABLES (USING THE CORRELATION MATRIX). THE REGRESSION ANALYSIS IS PERFORMED IN A STEPWISE MANNER AND THE RESULTING COEFFICIENTS ARE REPORTED IN TERMS OF BOTH PRINCIPAL COMPONENTS AND THE ORIGINAL OR STANDARDIZED VARIABLES. THE ORDER OF ENTRY OF COMPONENTS CAN BE BASED ON THE MAGNITUDE OF EIGENVALUES OR ON THE ABSOLUTE MAGNITUDE OF CORRELATION BETWEEN THE COMPONENT AND THE DEPENDENT VARIABLE. OUTPUT INCLUDES MEANS, STANDARD DEVIATIONS, THE COVARIANCE MATRIX (THE CORRELATION MATRIX IF STANDARDIZATION IS REQUESTED), EIGENVALUES AND EIGENVECTORS (PRINCIPAL COMPONENTS), REGRESSION COEFFICIENTS FOR PRINCIPAL COMPONENTS AND FOR INDEPENDENT VARIABLES, RESIDUAL SUM OF SQUARES AND F RATIOS FOR EACH STEP IN THE REGRESSION, AND F RATIOS FOR EACH COMPONENT. PRINCIPAL COMPONENT SCORES MAY BE REQUESTED, AS WELL AS SCATTER PLOTS OF RAW DATA, AND NORMAL AND DETRENDED NORMAL PROBABILITY PLOTS OF RESIDUALS.

BMDPSR — POLYNOMIAL REGRESSION
A POLYNOMIAL IN ONE VARIABLE IS FITTED TO THE DEPENDENT VARIABLE. COMPUTATION IS DONE USING ORTHOGONAL POLYNOMIALS. THE PROGRAM REPORTS POLYNOMIALS OF DEGREE ONE THROUGH A DEGREE SPECIFIED BY THE USER (LESS THAN OR EQUAL TO 15). ESTIMATES OF THE REGRESSION COEFFICIENTS AND THEIR STANDARD ERRORS ARE REPORTED. A SUMMARY TABLE OF GOODNESS-OF-FIT STATISTICS FOR EACH (EXCEPT THE HIGHEST DEGREE) POLYNOMIAL IS PRINTED. FOR EACH DEGREE A SCATTER PLOT OF THE PREDICTED AND OBSERVED VALUES VERSUS THE INDEPENDENT VARIABLE, AND RESIDUAL VERSUS THE INDEPENDENT VARIABLE MAY BE REQUESTED. THE USER MAY ALSO REQUEST NORMAL AND DETRENDED NORMAL PROBABILITY PLOTS OF THE RESIDUALS. THE RESIDUALS LISTED WITH THE DATA AND PREDICTED VALUES, AND A CORRELATION MATRIX FOR THE REGRESSION COEFFICIENT ESTIMATES.

BMDP6R — PARTIAL CORRELATION AND MULTIVARIATE REGRESSION
COMPUTES THE PARTIAL CORRELATIONS OF A SET OF VARIABLES REMOVING THE LINEAR EFFECTS OF A SECOND SET OF VARIABLES. THE PROGRAM CAN ALSO BE USED FOR REGRESSION, ESPECIALLY WHEN THERE ARE MULTIPLE DEPENDENT VARIABLES (SINCE THE COMPUTATIONS FOR PARTIAL CORRELATIONS INCLUDE THE COMPUTATIONS OF THE REGRESSION COEFFICIENTS FOR PREDICTING THE FIRST SET OF VARIABLES FROM THE SECOND SET OF VARIABLES). USING DOUBLE PRECISION, P6R CAN BE USED TO CHECK COMPUTATIONS FROM P1R AND P2R FOR PROBLEMS THAT MAY BE ILL-CONDITIONED OR INVOLVE A LARGE NUMBER (THOUSANDS) OF CASES. INPUT CAN BE A DATA MATRIX, A COVARIANCE MATRIX OR A CORRELATION MATRIX. OUTPUT INCLUDES MEANS, STANDARD DEVIATIONS, COEFFICIENTS OF VARIATION, SKEWNESS, KURTOSIS, MINIMUMS, MAXIMUMS, CORRELATIONS, COVARIANCES, PARTIAL CORRELATIONS, PARTIAL COVARIANCES, REGRESSION COEFFICIENTS, STANDARDIZED COEFFICIENTS, STANDARD ERRORS FOR COEFFICIENTS, COVARIANCES AND CORRELATIONS FOR REGRESSION COEFFICIENTS, T TESTS AND SIGNIFICANCE LEVELS FOR REGRESSION COEFFICIENTS, TWO TYPES OF SQUARED MULTIPLE CORRELATIONS AND THEIR SIGNIFICANCE LEVELS (THE FIRST, OF EACH INDEPENDENT VARIABLE WITH THE OTHER INDEPENDENT VARIABLES; THE SECOND, OF EACH DEPENDENT VARIABLE WITH ALL OF THE INDEPENDENT VARIABLES), RESIDUALS FROM PARTIAL CORRELATION ANALYSIS AND BIVARIATE PLOTS FOR ORIGINAL VARIABLES AND RESIDUALS.

BMDP9R ALL POSSIBLE SUBSETS REGRESSION
IDENTIFIES "BEST" SUBSETS OF PREDICTOR VARIABLES. BEST IS DEFINED IN TERMS OF THE SAMPLE R-SQUARE, ADJUSTED R-SQUARE, OR MALLOWS' C(P). FOR EXAMPLE, IF ADJUSTED R-SQUARE IS CHOSEN, THE BEST SUBSET IS THE SUBSET THAT MAXIMIZES ADJUSTED R-SQUARE. NOTE THAT BEST IS DEFINED AS NUMERICALLY BEST IN TERMS OF THE OBSERVED SAMPLE AND THAT (AS FOR STEPWISE REGRESSION) NO IMPLICATION IS MADE THAT ANY SUBSET CAN BE CONSIDERED BEST FOR THE POPULATION FROM WHICH THE SAMPLE WAS TAKEN. HOWEVER, (AS FOR STEPWISE REGRESSION) A WELL-CHOSEN SUBSET FOR ONE SAMPLE IS LIKELY TO BE GOOD FOR SIMILAR SAMPLES. THE NUMBER OF BEST SUBSETS CAN BE SPECIFIED; UP TO TEN CAN BE REQUESTED. THUS, NOT ONLY THE BEST BUT ALSO THE SECOND BEST, THIRD BEST, ETC. SUBSETS ARE IDENTIFIED TO PROVIDE SEVERAL GOOD ALTERNATIVES. WHEN M BEST SUBSETS ARE REQUESTED, THE R-SQUARE CRITERION REQUESTS THAT THE M BEST SUBSETS OF EACH SUBSET SIZE BE DETERMINED. WHEN THE ADJUSTED R-SQUARE OR MALLOWS' C(P) CRITERION IS CHOSEN, THE M BEST SUBSETS ARE FOUND REGARDLESS OF SUBSET SIZE. THE ALGORITHM USED IDENTIFIES THE BEST SUBSETS WHILE COMPUTING ONLY A SMALL FRACTION OF ALL POSSIBLE REGRESSIONS AND IS SUBSTANTIALLY MORE EFFICIENT THAN THOSE PREVIOUSLY AVAILABLE INPUT CAN BE DATA OR A CORRELATION OR COVARIANCE MATRIX. FOR UP TO APPROXIMATELY 27 VARIABLES, COMPUTER COSTS ARE COMPARABLE TO COSTS FOR STEPWISE REGRESSION. OUTPUT INCLUDES MEANS, STANDARD DEVIATIONS, COEFFICIENTS OF VARIATION, MINIMUMS, MAXIMUMS, MINIMUM AND MAXIMUM STANDARD SCORES, SKEWNESS AND KURTOSIS, CORRELATION MATRIX, COVARIANCE MATRIX, AND BIVARIATE PLOTS FOR ANY PAIR OF VARIABLES. FOR EACH OF THE M BEST SUBSETS, THE SELECTION CRITERION, REGRESSION COEFFICIENTS AND F RATIO ARE PRINTED. FOR THE BEST SUBSET P9R PRINTS R-SQUARE, F RATIO AND SIGNIFICANCE; REGRESSION COEFFICIENTS, STANDARDIZED REGRESSION COEFFICIENTS, STANDARD ERRORS, T TESTS, SIGNIFICANCE AND TOLERANCE; PREDICTED VALUES, RESIDUALS, WEIGHTED RESIDUALS, STANDARDIZED RESIDUALS, DELETED RESIDUALS, AND COOK'S MEASURE OF THE INFLUENCE OF EACH CASE ON THE REGRESSION EQUATION. BIVARIATE SCATTER PLOTS CAN BE REQUESTED FOR ANY PAIR OF VARIABLES INCLUDING RESIDUALS, STANDARDIZED RESIDUALS, ETC. A NORMAL PROBABILITY PLOT FOR THE STANDARDIZED RESIDUALS CAN BE PRINTED.

BMD01R SIMPLE LINEAR REGRESSION
(ONE-WAY ANALYSIS OF COVARIANCE)
SIMPLE LINEAR REGRESSION ANALYSIS ON SINGLE OR COMBINED CATEGORIES WITH UNEQUAL SAMPLE SIZES. ANALYSIS OF COVARIANCE INFORMATION IS ALSO OUTPUT. WILL HANDLE UP TO 999 GROUPS WITH UP TO 999 SUBJECTS.

BMD02R STEPWISE REGRESSION
COMPUTES A SEQUENCE OF MULTIPLE LINEAR REGRESSION EQUATIONS IN A STEPWISE MANNER. AT EACH STEP ONE VARIABLE IS ADDED TO THE REGRESSION EQUATION. THE VARIABLE ADDED IS THE ONE WHICH MAKES THE GREATEST REDUCTION IN THE ERROR SUM OF SQUARES. ALSO, VARIABLES CAN BE FORCED INTO THE REGRESSION EQUATION. NON-FORCED VARIABLES ARE AUTOMATICALLY REMOVED WHEN THEIR F VALUES BECOME TOO LOW. PLOTS OF RESIDUALS ARE AVAILABLE. WILL HANDLE 80 VARIABLES AND 9,999 SUBJECTS.

BMD03R MULTIPLE REGRESSION WITH CASE COMBINATIONS
PERFORMS MULTIPLE REGRESSION AND CORRELATION ANALYSES ON THE DATA WITHIN EACH SELECTION OF SUBSAMPLES FROM THE SAME POPULATION. SELECTIONS NEED NOT NOT BE MUTUALLY EXCLUSIVE. WILL HANDLE UP TO 28 SUBSAMPLES, 50 VARIABLES, AND TOTAL SAMPLE SIZE OF 99,999 FOR ALL SUBSAMPLES COMBINED.

BMD04R PERIODIC REGRESSION AND HARMONIC ANALYSIS
PERFORMS PERIODIC OR HARMONIC REGRESSION ANALYSIS USING THE MODEL,

$$Y(T) = A(0) + \text{THE SUM (FROM I = 1 TO N) OF } [A(I)\cdot\cos(2\cdot PI\cdot I\cdot T/K) + B(I)\cdot\sin(2\cdot PI\cdot I\cdot T/K)] ,$$

WHERE N IS LESS THAN OR EQUAL TO 9. PERIODIC REGRESSIONS ARE COMPUTED FOR SUCCESSIVE VALUES OF N, INCLUDING ANALYSIS OF VARIANCE AND/OR COVARIANCE TABLE FOR EACH REGRESSION.

BMD05R POLYNOMIAL REGRESSION
COMPUTES POLYNOMIAL REGRESSIONS OF THE FORM

$$Y = B(0) + B(1)\cdot X + B(2)\cdot X^{**}2 + \ldots + B(K)\cdot X^{**}K + E,$$

WHERE K IS A POSITIVE INTEGER LESS THAN OR EQUAL TO 10, AND THE NUMBER OF CASES IS LESS THAN OR EQUAL TO 500.

BMD06R — ASYMPTOTIC REGRESSION
PERFORMS ASYMPTOTIC REGRESSION ANALYSES USING THE MODIFIED EXPONENTIAL FUNCTION OF THE FORM

$$Y = A + B \cdot R^{**}X.$$

TRANSGENERATION OF Y CAN BE USED.

BMDX85 (OR BMD07R) — NONLINEAR LEAST SQUARES
THIS PROGRAM OBTAINS A LEAST-SQUARES FIT

$$Y(K) = F[X(1,K),\ldots,X(T,K);S(1),\ldots,S(P)] + E(K)$$

OF A USER-SPECIFIED FUNCTION F TO DATA VALUES X(I,K) AND Y(K) BY MEANS OF STEPWISE GAUSS-NEWTON ITERATIONS ON THE PARAMETERS S(1),...,S(P). IF NECESSARY, ONLY A SUBSET OF THE PARAMETERS WILL BE MODIFIED IN A GIVEN ITERATION IN ORDER TO AVOID THE PROBLEMS OF SINGULARITY.

CLASS S - SPECIAL PROGRAMS

BMDP1S — MULTIPASS TRANSFORMATION
P1S CAN BE USED WHEN INFORMATION FROM THE DATA FILE IS NEEDED TO COMPUTE TRANSFORMATIONS (I.E., ALL THE DATA MUST BE SCANNED OR PASSED BEFORE TRANSFORMAIONS CAN BE MADE; FOR EXAMPLE, MISSING VALUES MAY NEED TO BE REPLACED BY THE MEAN OF THE OF THE VALUES PRESENT IN THE SAMPLE). IN EACH PASS A FURTHER TRANSFORMATION OF THE DATA CAN BE SPECIFIED THROUGH FORTRAN STATEMENTS. ARITHMETIC MEANS, STANDARD DEVIATIONS, GEOMETRIC MEANS, HARMONIC MEANS, AND LARGEST AND SMALLEST VALUES FOR EACH VARIABLE CAN BE COMPUTED IN EACH TRANSFORMATION PASS TO BE USED IN THE NEXT PASS. (ALL THE BMDP PROGRAMS HAVE THE CAPABILITY OF DOING SIMPLE TRANSFORMATIONS AND DATA EDITING PROCEDURES.) OUTPUT FOR EACH PASS INCLUDES THE TRANSFORMED DATA AND ANY OF THE STATISTICS COMPUTED FOR EACH SELECTED VARIABLE. THE TRANSFORMED DATA CAN BE SAVED ON A BMDP FILE.

BMDP3S — NONPARAMETRIC STATISTICS
COMPUTES ONE OR MORE OF THE FOLLOWING NONPARAMETRIC STATISTICS FROM A GIVEN SET OF DATA: SIGN TEST, WILCOXON-SIGNED RANK TEST, KENDALL RANK CORRELATION COEFFICIENT, SPEARMAN RANK CORRELATION COEFFICIENT, FRIEDMAN TWO-WAY ANALYSIS OF VARIANCE, KENDALL

COEFFICIENT OF CONCORDANCE, MANN-WHITNEY U RANK SUM TEST, AND KRUSKAL-WALLIS ONE-WAY ANALYSIS OF VARIANCE.

BMD01S — LIFE TABLE AND SURVIVAL RATE
COMPUTES PROPORTIONS SURVIVING, SURVIVAL RATES, AND STANDARD ERRORS FOR SUCCESSIVELY REDUCED TIME PERIODS AND PRINTS A PLOT OF THE CUMULATIVE PROPORTION SURVIVING.

BMD02S — CONTINGENCY TABLE ANALYSIS
COMPUTES TWO-WAY FREQUENCY AND PERCENTAGE TABLES, CHI-SQUARES, CONTINGENCY COEFFICIENTS, AND MAXIMUM LIKELIHOOD RATIOS. EACH VARIABLE MAY BE CATEGORIZED IN SEVERAL DIFFERENT WAYS. OUTPUT INCLUDES FREQUENCY TABLES WITH PERCENTAGES COMPUTED FOR EACH ROW AND COLUMN. PERCENTAGES ARE ALSO COMPUTED FOR THE ENTIRE TABLE.

BMD03S — BIOLOGICAL ASSAY: PROBIT ANALYSIS
OBTAINS MAXIMUM LIKELIHOOD ESTIMATES FOR THE PARAMETERS A AND B IN THE PROBIT EQUATION

$$Y = A + B \cdot X.$$

AS AN OPTION, AN ESTIMATE OF THE NATURAL (THRESHOLD) RESPONSE RATE IS ALSO OBTAINED.

BMD09S — TRANSGENERATION
PERFORMS SELECTED TRANSGENERATIONS ON SPECIFIED VARIABLES IN THE DATA. ANY OF THE STANDARD BMD TRANSGENERATION CODES MAY BE SELECTED. SEE ALSO BMDX77 (OR BMD12S).

BMDX77 (OR BMD12S) — OPEN-ENDED TRANSGENERATION
PERFORMS TRANSGENERATION BY MEANS OF USER-SUPPLIED FORTRAN STATEMENTS. FOR STANDARD USE ONLY THE ARITHMETIC STATEMENTS ARE REQUIRED, ALTHOUGH THE ENTIRE FORTRAN IV LANGUAGE IS AVAILABLE. SEE ALSO BMD09S.

CLASS T - TIME SERIES ANALYSIS

BMD01T — AMPLITUDE AND PHASE ANALYSIS
COMPUTES AMPLITUDE AND PHASE OF MODERATELY WIDE-BAND NOISE AND NOISE CONTAMINATED BY EXTRANEOUS NOISE, USING A PAIR OF FINITE MOVING AVERAGES ON THE SAMPLE NOISE. A GENERALIZED TUKEY FILTER WITH

A VARIABLE NUMBER OF TRIANGLES AND RESOLUTIONS IS USED.

BMD02T — AUTOCOVARIANCE AND POWER SPECTRAL ANALYSIS
COMPUTES THE AUTOCOVARIANCE, POWER SPECTRUM, CROSS-COVARIANCE, CROSS-SPECTRUM, TRANSFER-FUNCTION, AND COHERENCE-FUNCTION OF TIME SERIES. PLOTS OF EACH ARE INCLUDED.

BMDX68 (OR BMD04T) — MULTIPLE TIME SERIES SPECTRAL ANALYSIS
BEGINNING WITH A SEQUENCE OF CROSS-SPECTRAL MATRICES, NORMALLY GENERATED BY BMDX92, THIS PROGRAM ESTIMATES MULTIPLE COHERENCE FUNCTIONS AND FREQUENCY-RESPONSE FUNCTIONS BETWEEN A SET OF INPUT AND A SET OF OUTPUT SERIES. CONFIDENCE BANDS ARE OBTAINED FOR THE GAIN AND PHASE OF THE COMPONENTS OF THE FREQUENCY RESPONSE FUNCTIONS. ESTIMATES OF THE CROSS-SPECTRA OF THE OUTPUT SERIES PARTIALED ON THE INPUT SERIES ARE ALSO OBTAINED.

BMDX92 (OR BMD03T) — TIME SERIES SPECTRUM ESTIMATION
ESTIMATES AUTO-SPECTRA, CROSS-SPECTRA, AND COHERENCES FOR STATIONARY TIME SERIES. EACH SERIES IS DECOMPOSED INTO ITS COMPONENT FREQUENCIES BY EVALUATING ITS FINITE FOURIER TRANSFORM.

CLASS V - VARIANCE ANALYSIS

BMDP1V — ONE-WAY ANALYSIS OF VARIANCE AND COVARIANCE
PERFORMS A ONE-WAY ANALYSIS OF VARIANCE, OR A ONE-WAY ANALYSIS OF COVARIANCE IF COVARIATES ARE SPECIFIED. GROUP SIZES CAN BE UNEQUAL AND PARALLEL ANALYSES CAN BE PERFORMED USING SEVERAL DEPENDENT VARIABLES. THE PROGRAM TESTS FOR EQUALITY OF SLOPES FROM GROUP TO GROUP AND COMPUTES T STATISTICS FOR CONTRASTS OF ADJUSTED GROUP MEANS INCLUDING, IN PARTICULAR, ALL PAIRWISE DIFFERENCES. THE ANALYSIS OF COVARIANCE IS ACCOMPANIED BY FOUR TYPES OF GRAPHS FOR EACH GROUP: SCATTER PLOTS OF EACH COVARIATE VERSUS THE DEPENDENT VARIABLE AND ITS PREDICTED VALUES, EACH COVARIATE VERSUS THE RESIDUALS, THE RESIDUAL VERSUS THE PREDICTED VALUES, AND THE SQUARED RESIDUALS VERSUS THE PREDICTED VALUES. OUTPUT FROM THE ANALYSIS OF VARIANCE ON EACH VARIABLE INCLUDES THE GROUP MEANS, AN ANALYSIS OF VARIANCE TABLE, PAIRWISE T TESTS FOR GROUP MEANS AND T TESTS FOR

CONTRASTS OF GROUP MEANS. FOR ANALYSIS OF COVARIANCE ON EACH DEPENDENT VARIABLE P1V REPORTS REGRESSION COEFFICIENTS, THEIR STANDARD ERRORS AND T VALUES, GROUP MEANS, ADJUSTED GROUP MEANS, STANDARD ERRORS OF ADJUSTED GROUP MEANS, AN ANALYSIS OF VARIANCE TABLE WITH F TEST FOR EQUALITY OF ADJUSTED GROUP MEANS, ZERO SLOPE, EQUALITY OF SLOPES, PAIRWISE T TEST FOR ADJUSTED GROUP MEANS, T TESTS FOR CONTRASTS OF ADJUSTED GROUP MEANS, AND REGRESSION COEFFICIENTS FOR EACH COVARIATE IN EACH GROUP. OTHER OUTPUT AVAILABLE INCLUDES THE CORRELATION MATRIX FOR THE REGRESSION COEFFICIENTS AND THE ADJUSTED GROUP MEANS, MAXIMUM, MINIMUM AND MEAN OF EACH VARIABLE IN EACH GROUP, VARIANCE-COVARIANCE MATRIX, CORRELATION MATRIX FOR EACH GROUP, AND TOTAL, BETWEEN AND WITHIN VARIANCE-COVARIANCE MATRICES.

BMDP2V — ANALYSIS OF VARIANCE AND COVARIANCE, INCLUDING REPEATED MEASURES

PERFORMS ANALYSES OF VARIANCE OR COVARIANCE FOR GENERAL FIXED EFFECTS AND REPEATED MEASURES MODELS. FOR EACH SUBJECT IN REPEATED MEASURES MODELS, THE TRIAL FACTORS (REPEATED MEASURES FACTORS, FOR WHICH EACH SUBJECT IS MEASURED AT EACH LEVEL) MUST HAVE A COMPLETE FACTORIAL STRUCTURE WITH NO MISSING OBSERVATIONS. GROUP INDICES ARE READ AS DATA AND DETERMINE THE GROUP MEMBERSHIPS WITHOUT FURTHER SPECIFICATION. P2V ALSO HANDLES MODELS THAT CONTAIN TRIAL FACTORS ONLY OR GROUPING FACTORS ONLY. WITHIN-SUBJECT RESPONSES NEED NOT BE INDEPENDENT, BUT BETWEEN-SUBJECT RESPONSES ARE ASSUMED TO BE. ALL FACTORS, EXCEPT SUBJECTS, ARE ASSUMED FIXED. AN ORTHOGONAL DECOMPOSITION OF THE TRIAL EFFECTS CAN BE REQUESTED. COVARIATES MAY BE (BUT NEED NOT BE) CONSTANT ACROSS TRIALS. UNEQUAL CELL SIZES ARE ALLOWED. SEVERAL DISTINCT HYPOTHESES CAN BE TESTED, DEPENDING ON THE PROBLEM. THE HYPOTHESES TESTED IN THIS PROGRAM ARE GENERALLY APPROPRIATE FOR EXPERIMENTAL DATA. OUTPUT INCLUDES AN ANALYSIS OF VARIANCE TABLE CONSISTING OF SUMS OF SQUARES, DEGREES OF FREEDOM, MEAN SQUARES, AND F STATISTICS (WITH PROBABILITY VALUES ASSOCIATED WITH EACH), CELL MEANS AND STANDARD DEVIATIONS FOR THE DEPENDENT VARIABLE AND COVARIATES, ADJUSTED CELL MEANS AND REGRESSION COEFFICIENTS (WHEN COVARIATES ARE PRESENT), AND DESIGN INFORMATION.

BMD01V — ANALYSIS OF VARIANCE OF ONE-WAY DESIGN
COMPUTES AN ANALYSIS OF VARIANCE TABLE FOR ONE VARIABLE OF CLASSIFICATION, WITH UNEQUAL GROUP SAMPLE SIZES. PRINT-OUT IS SELF-EXPLANATORY. THE PROGRAM TAKES UP TO 5,000 GROUPS, WITH A TOTAL SAMPLE SIZE OF 100,000,000.

BMD02V — ANALYSIS OF VARIANCE FOR FACTORIAL DESIGN
COMPUTES ANALYSIS OF VARIANCE FOR A FACTORIAL DESIGN. A BREAKDOWN OF THE SUMS OF SQUARES INTO ORTHOGONAL POLYNOMIAL COMPONENTS FOR AS MANY AS FOUR MAIN EFFECTS AND ALL OF THEIR FIRST ORDER INTERACTIONS FOR UP TO EIGHT SPECIFIED VARIABLES ARE COMPUTED. CELL AND MARGINAL MEANS ARE LISTED. PRINT-OUT IS SELF-EXPLANATORY. PROGRAM WILL TAKE UP TO EIGHT FACTORS WITH 999 SUBJECTS PER CELL, CELLS NOT TO EXCEED 18,000 IN NUMBER. WILL ALSO DO TREND ANALYSIS.

BMD03V — ANALYSIS OF COVARIANCE FOR FACTORIAL DESIGN
COMPUTES A FULL FACTORIAL ANALYSIS OF COVARIANCE FOR A MAXIMUM OF SIX FACTORIAL COMPONENTS AND A MAXIMUM OF EIGHT COVARIATES. ADJUSTED RESIDUALS AND REGRESSION COEFFICIENTS ARE COMPUTED. THE INVERSE OF THE COVARIANCE MATRIX OF THE COVARIATES IS PRINTED.

BMD04V — ANALYSIS OF COVARIANCE WITH MULTIPLE COVARIATES
COMPUTES ANALYSIS OF COVARIANCE INFORMATION FOR ONE ANALYSIS OF VARIANCE VARIABLE WITH MULTIPLE COVARIATES AND UNEQUAL TREATMENT GROUP SIZES. CASES MAY BE SPECIFIED BY THE USER AS BEING IN CERTAIN CATEGORIES, OR CASES MAY BE PLACED IN CATEGORIES BY THE PROGRAM IN ACCORDANCE WITH A SPECIFIED CRITERION. WILL TAKE UP TO 35 COVARIATES IN 99 GROUPS, NONE TO EXCEED 999 SUBJECTS. SEE ALSO BMDX82 AND BMD01R.

BMD05V — GENERAL LINEAR HYPOTHESIS
PERFORMS THE CALCULATIONS REQUIRED FOR A GENERAL LINEAR HYPOTHESIS MODEL. CAN BE USED FOR BALANCED OR UNBALANCED ANALYSIS OF VARIANCE OR COVARIANCE DESIGNS AND MISSING-VALUE PROBLEMS. THE INDEPENDENT VARIABLES ARE OF TWO GENERAL TYPES: THE VARIABLES SPECIFYING THE ANALYSIS OF VARIANCE CLASSIFICATIONS AND THE VARIABLES WHICH ARE COVARIATES. SEE ALSO BMDX69 (OR BMD10V).

BMD06V — GENERAL LINEAR HYPOTHESIS WITH CONTRASTS
SIMILAR TO BMD05V: DESIGNED TO ESTIMATE AND TEST

THE STATISTICAL SIGNIFICANCE OF THE PARAMETERS WHICH OCCUR IN THE GENERAL LINEAR HYPOTHESIS MODEL. ALSO, IT CAN TEST THE STATISTICAL SIGNIFICANCE OF ANY REAL-VALUED LINEAR FUNCTION OF THE PARAMETERS. HIGHER-ORDER SIMULTANEOUS TESTS CANNOT BE MADE. ALSO SEE BMDX69 (OR BMD10V).

BMD07V — MULTIPLE RANGE TESTS
PERFORMS MULTIPLE RANGE TESTS ON TWO KINDS OF INPUT: (1) RAW DATA WITH ONE VARIABLE OF CLASSIFICATION AND UNEQUAL GROUP SIZES AND (2) SAMPLE SIZES, GROUP MEANS, STANDARD DEVIATIONS, AND GROUP LABELS. TEST IS DONE ACCORDING TO USER CARD SPECIFICATIONS; OTHERWISE IT WILL BE DONE USING DUNCAN'S MULTIPLE RANGE TEST SUPPLIED BY THE PROGRAM.

BMD08V — ANALYSIS OF VARIANCE
PERFORMS ANALYSIS OF VARIANCE FOR ANY HIERARCHICAL DESIGN WITH EQUAL CELL SIZES. THIS INCLUDES NESTED, PARTIALLY NESTED AND PARTIALLY CROSSED, AND FULLY CROSSED DESIGNS. THE MODEL IS SPECIFIED BY INDICATING THE NESTING RELATIONSHIPS OF THE INDICES. SEPARATE ANALYSES MAY BE PERFORMED SIMULTANEOUSLY ON SEVERAL DEPENDENT VARIABLES. EACH ANALYSIS OF VARIANCE TABLE INCLUDES AN EXPECTED MEAN SQUARE IN TERMS OF THE POPULATION VARIANCE COMPONENTS. VERY VERSATILE AS COMPARED WITH BMD01V AND BMD02V.

BMDX63 (OR BMD11V) — MULTIVARIATE GENERAL LINEAR HYPOTHESIS
PERFORMS MULTIPLE REGRESSION ON A DEPENDENT VARIABLE IN THE FORM OF A VECTOR. COMPUTES LIKELIHOOD RATIO STATISTICS AND APPROXIMATE F-STATISTIC TO TEST MULTIVARIATE HYPOTHESES OF THE FORM $A\ B\ C' = D$ WHERE B IS A MATRIX OF REGRESSION COEFFICIENTS AND A, C, AND D ARE MATRICES SPECIFIED BY THE USER. WITH PROPER SPECIFICATION, THE PROGRAM CAN BE USED TO CARRY OUT BALANCED OR UNBALANCED MULTIVARIATE ANALYSES OF VARIANCE AND COVARIANCE.

BMDX69 (OR BMD10V) — GENERAL LINEAR HYPOTHESIS
ESTIMATES PARAMETERS AND TESTS HYPOTHESES CONCERNING A GENERAL LINEAR MODEL. DESIGN VARIABLES FOR THE ANALYSIS OF VARIANCE ARE GENERATED BY THE PROGRAM. MODEL I ANALYSES OF VARIANCE AND COVARIANCE TESTS ARE PERFORMED AUTOMATICALLY ALONG WITH TESTS OF USER-SPECIFIED LINEAR HYPOTHESES. THE PROGRAM DOES NOT REQUIRE BALANCED OR EVEN FULL RANK MODELS.

BMDX69 (OR BMD12V)
MULTIVARIATE ANALYSIS OF VARIANCE AND COVARIANCE
PERFORMS MODEL I UNIVARIATE OR MULTIVARIATE ANALYSIS OF VARIANCE OR COVARIANCE FOR ANY COMPLETE BALANCED DESIGN WITH EQUAL CELL SIZES. THE DESIGN IS SPECIFIED BY INDICATING NESTING RELATIONS FOR THE CLASSIFICATION INDICES.

BMDX82 (OR BMDC9V)
ANALYSIS OF COVARIANCE
PERFORMS A ONE-WAY ANALYSIS OF COVARIANCE, USING ONE OR MORE COVARIATES. GROUP SIZES MAY BE UNEQUAL AND PARALLEL ANALYSES MAY BE PERFORMED USING SEVERAL DEPENDENT VARIABLES. THE PROGRAM TESTS FOR EQUALITY OF SLOPE FROM GROUP TO GROUP AND COMPUTES T-STATISTICS FOR CONTRASTS OF ADJUSTED GROUP MEANS, INCLUDING, IN PARTICULAR, ALL PAIRWISE DIFFERENCES.

9.1.3. DOCUMENTATION REFERENCES

DIXON, W.J. (ED.). BMD BIOMEDICAL COMPUTER PROGRAMS. BERKELEY, CALIFORNIA: UNIVERSITY OF CALIFORNIA PRESS. 1968.

DIXON, W.J. (ED.). BMD BIOMEDICAL COMPUTER PROGRAMS P-SERIES SUPPLEMENT. BERKELEY, CALIFORNIA: UNIVERSITY OF CALIFORNIA PRESS. 1970.

DIXON, W.J. (ED.). BMD BIOMEDICAL COMPUTER PROGRAMS. BERKELEY, CALIFORNIA: UNIVERSITY OF CALIFORNIA PRESS. 1973.

DIXON, W.J., & BROWN, M.B. (EDITORS). BMDP-77 BIOMEDICAL COMPUTER PROGRAMS, P-SERIES. BERKELEY, CALIFORNIA: UNIVERSITY OF CALIFORNIA PRESS. 1977.

Mathematical and Statistical Software Index

Introduction 1

Mathematical and Statistical Software Index 3

This index is a catalog of computer programs available from the Statistical Research Laboratory. Each program is described briefly, and the name of the author is given. Programs are grouped by subject matter. Within each group, programs are listed in alphabetical order by author's last name. Programs written by more than one author are listed under the name of the author who first submitted the program to the Laboratory.

Each program is described briefly. The description includes the purpose of the program, the method used, and any restrictions on its use. Programs are written in FORTRAN IV unless otherwise noted. Machine requirements are given for each program. Programs are available on magnetic tape or punched cards. Requests should be addressed to the Laboratory.

[illegible]

[illegible]

9.2.2. LIST OF PROGRAMS AND PROGRAM DESCRIPTIONS

COLFREQ: SINGLE COLUMN FREQUENCY COUNTS
PERFORMS SINGLE COLUMN FREQUENCY COUNTS OF LETTERS, NUMERALS, AND OTHER CHARACTERS FOR SPECIFIED COLUMNS OF INPUT RECORDS UP TO 2000 CHARACTERS IN LENGTH. IN ADDITION TO A TABLE OF FREQUENCIES FOR THE CHARACTERS APPEARING IN SPECIFIED COLUMNS, STANDARD OUTPUT ALSO INCLUDES A SUMMARY, BY COLUMN, OF THE FREQUENCY COUNTS AND PERCENTAGES FOR FOUR GROUPS OF CHARACTERS: NUMERALS, LETTERS, BLANKS, AND SPECIAL CHARACTERS. ALSO INCLUDED ARE THE NUMERIC MEAN OF EACH COLUMN AND THE ADJUSTED MEAN, TREATING BLANKS AS ZEROES. AS OPTIONAL OUTPUT, A TABLE OF PERCENTAGES BY COLUMN IS AVAILABLE GIVING THE PERCENT OF FREQUENCY OF EACH CHARACTER.

CROSTAB2: DATA TABULATION
PRODUCES MULTIDIMENSIONAL CROSS-TABULATIONS WITH NO RESTRICTIONS ON THE NUMBER OF DIMENSIONS. IN EACH TABULATION, ONE OR MORE VARIABLES MAY BE TREATED AS CLASSIFICATION VARIABLES WHILE OTHER VARIABLES ARE USED AS ACCUMULATION VARIABLES TO OBTAIN CELL MEANS, STANDARD DEVIATIONS AND OTHER STATISTICS, IN ADDITION TO FREQUENCY COUNTS AND PERCENTS. EITHER NUMERIC (CONTINUOUS) VARIABLES OR ALPHANUMERIC (DISCRETE) VARIABLES, OR BOTH, CAN BE USED AS CLASSIFICATION VARIABLES. THERE ARE FACILITIES FOR SPECIFICATION OF CATEGORIES OF VALUES, AND LABELS AND WEIGHTS ASSOCIATED WITH THEM. DISTINCTION CAN BE MADE FOR CLASSIFICATION VARIABLES OF PRIMARY AND SECONDARY INTEREST. MARGINAL TOTALS, PERCENTAGES, AND BIVARIATE STATISTICS ARE GIVEN BASED ON PRIMARY CATEGORIES ONLY. CHI-SQUARE, TAU, AND GAMMA BIVARIATE STATISTICS CAN BE COMPUTED OR A USER-SUPPLIED ROUTINE CAN COMPUTE BIVARIATE STATISTICS. SPECIAL PROCEDURES ARE INCLUDED FOR HANDLING MISSING DATA.

DISCRIM: DISCRIMINANT ANALYSIS
PERFORMS DISCRIMINANT ANALYSIS ON TWO OR MORE GROUPS OF DATA USING A SET OF PREDICTOR OR DISCRIMINATING VARIABLES. CLASSICAL AND STEPWISE VARIABLE SELECTION PROCEDURES, OR A COMBINATION OF THESE PROCEDURES, ARE PROVIDED WITH RAO'S V STATISTIC BEING USED AS THE CRITERION FOR ENTERING A VARIABLE OR A SET OF VARIABLES INTO THE ANALYSIS. OUTPUT INCLUDES DESCRIPTIVE STATISTICS BY GROUP AND FOR THE TOTAL SAMPLE: BIVAR-

IATE STATISTICS (MOMENTS, COVARIANCES, CORRELATIONS) FOR EACH GROUP, WITHIN GROUPS, AMONG GROUPS, AND FOR THE TOTAL SAMPLE; AND MAHALANOBIS' D-SQUARE STATISTIC FOR PAIRS OF GROUPS. A UNIVARIATE ANALYSIS OF VARIANCE ON EACH VARIABLE MAY ALSO BE PROVIDED, AS WELL AS A TEST OF THE EQUALITY OF GROUP COVARIANCE MATRICES. FOR EACH STEP IN A STEPWISE ANALYSIS, RAO'S V STATISTIC AND SIGNIFICANCE LEVEL ARE PROVIDED. WHILE OUTPUT FOR THE FINAL SOLUTION OF ANALYSIS INCLUDES WILKS' LAMBDA TEST OF THE EQUALITY OF GROUP MEANS; EIGENVALUES AND EIGENVECTORS AND OTHER RESULTS ASSOCIATED WITH THE DISCRIMINANT FUNCTIONS; AND THE REDUCED SPACE CENTROIDS, DISPERSION MATRICES, AND ASSOCIATED RESULTS.

DSTAT2 DESCRIPTIVE STATISTICS AND CORRELATION
COMPUTES A CORRELATION MATRIX AND A SET OF DESCRIPTIVE STATISTICS FOR A DATA SET CONSISTING OF OBSERVATIONS ON A SET OF VARIABLES. RAW INPUT DATA MAY CONTAIN MISSING DATA. OPTIONALLY PROVIDES FISHER'S Z-TRANSFORMATION OF THE CORRELATION COEFFICIENTS AND CORRESPONDING SIGNIFICANCE LEVELS; T-TRANSFORMATIONS OF THE CORRELATIONS AND CORRESPONDING SIGNIFICANCE LEVELS; THE COVARIANCE MATRIX; AND THE RAW CROSS-PRODUCTS MATRIX. FOR MISSING DATA, THE CORRELATION MATRIX IS COMPUTED USING THE BIVARIATE SUBSAMPLE METHOD, OR OPTIONALLY, THE REPLACEMENT WITH MEANS METHOD.

FACTOR3 FACTOR ANALYSIS
COMPUTES A FACTOR LOADING OR PATTERN MATRIX EITHER BY FACTORING THE COVARIANCE MATRIX (FOR A SET OF INPUT VARIABLES) WITH THE PRINCIPAL COMPONENTS PROCEDURE OR BY FACTORING THE CORRELATION MATRIX OF THE INPUT VARIABLES USING ONE OF FIVE PROCEDURES: PRINCIPAL COMPONENTS; PRINCIPAL FACTOR; UNIQUENESS RESCALING OR, OPTIONALLY, RAO'S CANONICAL FACTOR PROCEDURE; ALPHA; AND IMAGE. COMMUNALITY ESTIMATES MAY BE PROVIDED OR THEY MAY BE DETERMINED FROM THE DATA AS SQUARED MULTIPLE CORRELATIONS. VARIMAX OR QUARTIMAX ROTATIONS ARE OPTIONALLY AVAILABLE. FACTOR SCORES MAY BE REQUESTED FOR THE PRINCIPAL COMPONENTS AND IMAGE PROCEDURES. STANDARD OUTPUT INCLUDES CORRELATIONS, DESCRIPTIVE STATISTICS, COMMUNALITIES, EIGENVALUES, FACTOR LOADINGS, AND ANALYSIS OF FACTOR VARIANCE.

MANAY1 GENERAL ANALYSIS OF VARIANCE
PERFORMS AN ANALYSIS OF VARIANCE FOR ANY COMPLETE FACTORIAL DESIGN WITH UP TO 12 FACTORS. EACH PAIR OF FACTORS MUST BE EITHER COMPLETELY CROSSED OR COMPLETELY NESTED. IF THE CELLS OF THE DESIGN DO NOT HAVE THE

SAME NUMBER OF REPLICATES, AND IF REPLICATES ARE NOT USED AS A FACTOR, THEN AN UNWEIGHTED MEANS ANALYSIS IS PERFORMED (EXCEPT FOR ONE-WAY ANALYSIS WHICH IS EXACT). IF REPLICATES ARE USED AS A FACTOR, THEN HENDERSON'S METHOD I IS USED. STANDARD OUTPUT FROM NWAY1 CONSISTS OF AN ANALYSIS OF VARIANCE TABLE CONTAINING, FOR EACH SOURCE OF VARIATION IN THE DESIGN, THE SUM OF SQUARES, DEGREES OF FREEDOM, MEAN SQUARE, F-RATIO AND SIGNIFICANCE LEVEL, AND ESTIMATE OF VARIANCE COMPONENT. OPTIONAL OUTPUT INCLUDES TABLES OF CELL COUNTS, MEANS AND STANDARD DEVIATIONS OR VARIANCES; TABLES OF EXPECTED MEAN SQUARES; AND TABLES OF DENOMINATOR MEAN SQUARES AND DEGREES OF FREEDOM.

ONEWAY2 ONE-WAY ANALYSIS OF VARIANCE
PERFORMS ONE-WAY ANALYSIS OF VARIANCE AND PERMITS COMPARISONS OF CELL MEANS BY T-TESTS, LINEAR CONTRASTS, ORTHOGONAL POLYNOMIAL CONTRASTS, MULTIPLE RANGE TESTS, AND SIMPLE-EFFECTS ANALYSIS OF VARIANCE. STANDARD OUTPUT CONSISTS OF SEVERAL DESCRIPTIVE STATISTICS FOR EACH GROUP AND THE ANALYSIS OF VARIANCE TABLE. OPTIONAL OUTPUT INCLUDES RESIDUALS, HISTOGRAMS FOR EACH GROUP, ETA SQUARED AND OMEGA SQUARED, ESTIMATE OF GROUP VARIANCES FOR THE RANDOM EFFECTS MODEL, AND OUTPUT FOR THE VARIOUS TYPES OF CELL MEAN COMPARISONS AVAILABLE. OPTIONAL OUTPUT IS ALSO GIVEN FOR THE FOLLOWING TESTS OF HOMOGENEITY OF VARIANCE: BARTLETT'S TEST; COCHRAN'S TEST; F-MAX TEST; BARTLETT-KENDALL-BOX TEST; BOX-ANDERSON TEST; LEVENE TEST; AND JACKKNIFE TEST. PAIRED T-TESTS OR ONE SAMPLE T-TESTS CAN BE PERFORMED BY CALCULATING DIFFERENCE SCORES AND EXAMINING THE TEST FOR THE MEAN IN THE ANALYSIS OF VARIANCE TABLE. UP TO 1000 VARIABLES CAN BE ANALYZED IN ONE COMPUTER RUN WITH THE MAXIMUM NUMBER OF GROUPS ALLOWED RANGING FROM 400 TO 1000 DEPENDING WHICH OPTIONAL OUTPUTS ARE REQUESTED.

PICT1 SCATTERPLOTS, GRAPHS, AND TABLES
PROVIDES BIVARIATE SCATTERPLOTS, GRAPHS AND TABLES OF THE VALUES OF SELECTED VARIABLES FROM A DATA SET. SCATTERPLOTS CONSIST OF PLOTS OF THE DATA VALUES OF PAIRS OF INPUT VARIABLES ON A TWO-DIMENSIONAL COORDINATE SYSTEM. FACILITIES ARE PROVIDED TO DENOTE UP TO THIRTY-FIVE OCCURRENCES OF A PARTICULAR COORDINATE POINT AND TO CLASSIFY DATA POINTS INTO GROUPS ON THE BASIS OF A THIRD VARIABLE WITH A UNIQUE SYMBOL FOR EACH GROUP. OPTIONAL OUTPUT WITH THE SCATTERPLOTS INCLUDES REGRESSION STATISTICS AND INDICATORS OF THE REGRESSION LINE ON THE PLOT AND A VARIETY OF GRID SYSTEMS. GRAPHS CONSIST OF VALUES FOR UP TO EIGHT DEPEN-

DENT VARIABLES PLOTTED ON THE SAME GRAPH USING THE SAME SCALE FOR ALL VARIABLES ALONG THE VERTICAL AXIS. THE HORIZONTAL AXIS REPRESENTS THE ORDER OF OBSERVATIONS OF THE DATA SET OR THE VALUES, SORTED IN ASCENDING ORDER, OF A SPECIFIED VARIABLE IN THE DATA SET. WHEN MORE THAN ONE VARIABLE IS INCLUDED IN THE VERTICAL DIRECTION OF THE GRAPH, THE CUMULATIVE SUMS OF SUCCESSIVE VARIABLES CAN BE PLOTTED INSTEAD OF THE VALUES THEMSELVES. TABLES PRODUCED WITH PICTI ARE PRODUCED BY LISTING THE VALUES OR CUMULATIVE SUMS OF UP TO EIGHT VARIABLES AGAINST THE ORDER OF THE OBSERVATIONS WITHIN THE DATA SET OR ACCORDING TO THE ASCENDING ORDER OF VALUES OF A SPECIFIED VARIABLE FROM THE DATA SET.

REGAN2 MULTIPLE LINEAR REGRESSION
PERFORMS MULTIPLE LINEAR REGRESSION ANALYSIS WITH A CAPABILITY OF SOLVING MANY MODELS IN ONE PASS OF THE DATA. INDEPENDENT VARIABLES MAY ENTER THE REGRESSION EQUATION IN GROUPS. STANDARD OUTPUT INCLUDES THE CORRELATION MATRIX AND DESCRIPTIVE STATISTICS FOR ALL INPUT VARIABLES. AS EACH GROUP OF VARIABLES ENTERS THE EQUATION, THE FOLLOWING OUTPUT MAY BE OBTAINED: ESTIMATE OF STANDARD ERROR OF THE REGRESSION; MULTIPLE CORRELATION COEFFICIENT R, R-SQUARE, AND CORRECTED R-SQUARE; ORDINARY AND STANDARDIZED REGRESSION COEFFICIENTS; STANDARD ERRORS OF THE REGRESSION COEFFICIENTS; PARTIAL CORRELATION COEFFICIENTS; PARTIAL F VALUE, T-VALUE, AND SIGNIFICANCE LEVEL FOR EACH VARIABLE; PARTIAL F VALUE AND SIGNIFICANCE LEVEL FOR EACH GROUP OF VARIABLES ENTERING; ANOVA TABLE; COVARIANCE OR CORRELATION MATRIX OF REGRESSION COEFFICIENTS; RESIDUALS AND PREDICTED VALUES, WITH A PLOT OF STANDARDIZED RESIDUALS.

ROTATE1 FACTOR ROTATION
PERFORMS FACTOR ANALYTIC ROTATIONS OF AN INPUT MATRIX OF FACTOR LOADINGS, USUALLY FROM AN UNROTATED FACTOR PATTERN MATRIX. HOWEVER, A ROTATED PATTERN MATRIX MAY ALSO BE INPUT. THE USER MAY SELECT COLUMNS FROM THE LOADING MATRIX FOR ROTATION. MORE THAN ONE TYPE OF ROTATION CAN BE MADE, AND MORE THAN ONE SET OF FACTORS CAN BE ROTATED IN A SINGLE RUN. BOTH ORTHOGONAL AND OBLIQUE ROTATIONS ARE AVAILABLE WITH THE ORTHOGONAL ROTATIONS FROM THE ORTHOMAX FAMILY, TWO CASES OF WHICH ARE VARIMAX AND QUARTIMAX ROTATION. TWO OBLIQUE ROTATIONS MAY BE SELECTED FROM THE HARRIS AND KAISER FAMILY OF "ORTHOBLIQUE" ROTATIONS: INDEPENDENT CLUSTER ROTATION (TYPE 1) AND P'P PROPORTIONAL TO FACTOR CORRELATION MATRIX (TYPE 2). STANDARD OUTPUT

INCLUDES RAW FACTOR MATRIX; ROTATED FACTOR PATTERN MATRIX; VARIANCE ANALYSIS FOR ROTATED FACTORS; FACTOR STRUCTURE MATRIX (OBLIQUE ROTATIONS ONLY); AND CORRELATION MATRIX FOR OBLIQUELY ROTATED FACTOR SCORES.

STEPREG1 STEPWISE LINEAR REGRESSION ANALYSIS
PERFORMS MULTIPLE LINEAR REGRESSION IN A STEPWISE MANNER WHICH IS AN EXPANSION OF THE PROCEDURE DESCRIBED BY EFROYMSON (PAGES 191-203) IN "MATHEMATICAL METHODS FOR DIGITAL COMPUTERS," BY R. RALSTON AND H. WILF (EDITORS). ONE SET OF INDEPENDENT VARIABLES (BASE VARIABLES) CAN BE FORCED INTO THE REGRESSION EQUATION IN THE ORDER SPECIFIED ON A CONTROL CARD; THEY MAY ENTER THE EQUATION EITHER INDIVIDUALLY, OR IN GROUPS. INDEPENDENT VARIABLES FROM ANOTHER SET (FREE VARIABLES) CAN BE SELECTED FOR ENTRY OR REMOVAL FROM THE EQUATION WITH EITHER FORWARD OR BACKWARD STEPWISE PROCEDURES. EACH PROCEDURE HAS OPTIONS TO SPECIFY SIGNIFICANCE CRITERIA FOR ENTRY OR REMOVAL OF INDEPENDENT VARIABLES. THE CORRELATION MATRIX AND DESCRIPTIVE STATISTICS ARE PRINTED FOR ALL INPUT VARIABLES. OUTPUT MAY BE OBTAINED AT EACH STEP, OR AT THE FINAL STEP IN THE SOLUTION PROCESS, AND IS IDENTICAL WITH THE OUTPUT THAT MAY BE OBTAINED WITH REGAN2. AFTER COMPLETION OF THE ANALYSIS, A SUMMARY OF STEPS IS GIVEN, FOLLOWED BY A LIST OF STATISTICS RELATED TO VARIABLES NOT IN THE FINAL EQUATION.

TRANS1 DATA TRANSFER
TRANS1 CAN BE USED TO SORT DATA SETS ON ONE OR MORE VARIABLES IN ASCENDING OR DESCENDING ORDER. ONE OR MORE VARIABLES MAY BE RANK ORDERED, PERCENTILE RANKED, OR STANDARDIZED. TRANS1 CAN BE USED TO STRATIFY A DATA SET INTO SEVERAL DATA SETS ON THE VALUE OF AN INPUT VARIABLE. TRANS1 PERFORMS TWO OTHER FUNCTIONS OF A ROUTINE DATA PROCESSING NATURE: TRANSFER OF DATA, WITH OR WITHOUT TRANSFORMATION, FROM ONE DATA MODE AND/OR MEDIA TO ANOTHER; AND SERVES AS THE INPUT ROUTINE, WITH OR WITHOUT TRANSFORMATION, TO A USER-SUPPLIED ANALYSIS ROUTINE.

UNISTAT2 UNIVARIATE STATISTICS AND HISTOGRAMS
PRINTS UNIVARIATE DESCRIPTIVE STATISTICS AND HISTOGRAMS FOR NUMERIC VARIABLES, AND HISTOGRAMS FOR ALPHANUMERIC VARIABLES. ONE OR MORE SUBSETS OF VARIABLES CAN BE SPECIFIED FOR ANALYSIS. A WEIGHT VARIABLE CAN BE SPECIFIED FOR ALL COMPUTATIONS, OR A SEPARATE WEIGHT VARIABLE CAN BE SPECIFIED WITH EACH SUBSET. THE DESCRIPTIVE STATISTICS INCLUDE MEAN, VARIANCE, STANDARD DEVIATION, MINIMUM AND MAXIMUM

VALUES, STANDARD ERROR OF THE MEAN, COEFFICIENT OF VARIATION, STUDENTIZED MEAN, MEAN ABSOLUTE DEVIATION, AND COEFFICIENTS OF SKEWNESS AND KURTOSIS; CONFIDENCE INTERVALS ON THE MEAN AND VARIANCE; SELECTED PERCENTILES, MEDIAN, AND TRI-MEAN; TRIMMED AND WINSORIZED MEAN; AUTOCORRELATIONS UP TO ANY ORDER; AND INFORMATION ON RUNS UP AND DOWN, AND RUNS ABOVE AND BELOW THE MEDIAN. FOR NUMERIC HISTOGRAMS, EXPECTED AS WELL AS OBSERVED FREQUENCIES CAN BE PLOTTED FOR THE NORMAL OR POISSON PROBABILITY MODELS. WHEN EXPECTED FREQUENCIES ARE PLOTTED, PEARSON'S CHI-SQUARE GOODNESS-OF-FIT TEST IS PERFORMED.

4.2.3. DOCUMENTATION REFERENCES

STATJOB SERIES OF REFERENCE MANUALS FOR UNIVAC 1100 SERIES COMPUTERS, ACADEMIC COMPUTING CENTER, UNIVERSITY OF WISCONSIN, MADISON, WISCONSIN.

THE SERIES INCLUDES THE FOLLOWING MANUALS:

INTRODUCTION TO STATJOB, VERSION 10
STATJOB SUMMARY, VERSION 10
COLFREQ1, SINGLE COLUMN FREQUENCY COUNTS, DECEMBER, 1971
CROSTAB2, DATA TABULATION, MARCH, 1975
DISCRIM1, DISCRIMINANT ANALYSIS, MARCH, 1975
DSTAT2, DESCRIPTIVE STATISTICS AND CORRELATION, SEPTEMBER, 1973
FACTOR3, FACTOR ANALYSIS, MAY, 1974
NWAY1 FOR PSYCHOLOGISTS, MAY, 1971
ONEWAY2, ONE-WAY ANALYSIS OF VARIANCE, NOVEMBER, 1975
PICT1, SCATTERPLOTS, GRAPHS, AND TABLES, SECOND REVISION, APRIL, 1976
REGAN2, MULTIPLE LINEAR REGRESSION ANALYSIS, APRIL, 1971
ROTATE1, FACTOR ROTATION, NOVEMBER, 1974
STEPREG1, STEPWISE REGRESSION ANALYSIS, AUGUST, 1973
TRANS1, DATA TRANSFER, FIRST REVISION, DECEMBER, 1974
UNISTAT2, UNIVARIATE DESCRIPTIVE STATISTICS AND HISTOGRAMS, APRIL, 1976
BASIC SYSTEM PROCEDURES, MAY, 1970
EXTENDED SYSTEM PROCEDURES, FEBRUARY, 1972
STJBANK PROCEDURES, OCTOBER, 1974
TRANSFORMATION PROCEDURES, OCTOBER, 1975.

REFERENCES

BAKER, F.B. ORIGINS OF THE ITEM PARAMETERS X50 AND BETA AS A MODERN ITEM ANALYSIS TECHNIQUE. JOURNAL OF EDUCATIONAL MEASUREMENT, 1965, 2, 167-178.

BOTTENBERG, R.A., & CHRISTAL, R.E. AN ITERATIVE TECHNIQUE FOR CLUSTERING CRITERIA WHICH RETAINS OPTIMUM PREDICTIVE EFFICIENCY. WADD-TN-61-30, AD-261 615. LACKLAND AFB, TX: PERSONNEL LABORATORY, WRIGHT AIR DEVELOPMENT DIVISION, MARCH 1961.

CHISMAN, J.A. AN INVESTIGATION INTO THE USE OF FOURIER ANALYSIS IN STATISTICAL CURVE FITTING. ENGINEERING EXPERIMENTAL STATION, BULLETIN NO. 112. COLLEGE OF ENGINEERING, CLEMSON UNIVERSITY, 1968. (A)

CHISMAN, J.A. THE PEARSON GENERALIZED STATISTICAL DISTRIBUTION. ENGINEERING EXPERIMENTAL STATION, BULLETIN NO. 111. COLLEGE OF ENGINEERING, CLEMSON UNIVERSITY, 1968. (B)

COOLEY, W.W., & LOHNES, P.R. MULTIVARIATE PROCEDURES FOR THE BEHAVIORAL SCIENCES. NEW YORK: JOHN WILEY AND SONS, INCORPORATED, 1962.

DIXON, W.J. (ED.). BMD BIOMEDICAL COMPUTER PROGRAMS. BERKELEY, CALIFORNIA: UNIVERSITY OF CALIFORNIA PRESS, 1968.

DIXON, W.J. (ED.). BMD BIOMEDICAL COMPUTER PROGRAMS X-SERIES SUPPLEMENT. BERKELEY, CALIFORNIA: UNIVERSITY OF CALIFORNIA PRESS, 1970.

DIXON, W.J. (ED.). BMD BIOMEDICAL COMPUTER PROGRAMS. BERKELEY, CALIFORNIA: UNIVERSITY OF CALIFORNIA PRESS, 1973.

DIXON, W.J., & BROWN, M.B. (EDITORS). BMDP-77 BIOMEDICAL COMPUTER PROGRAMS, P-SERIES. BERKELEY, CALIFORNIA: UNIVERSITY OF CALIFORNIA PRESS, 1977.

EFROYMSON, M.A. MULTIPLE REGRESSION ANALYSIS. IN R. RALSTON & H. WILF (EDS.), MATHEMATICAL METHODS FOR DIGITAL COMPUTERS (VOL. 1). NEW YORK: JOHN WILEY AND SONS, INCORPORATED, 1960.

ELDERTON, W.P. FREQUENCY CURVES AND CORRELATION. LONDON: HARREN PRESS, 1953.

FRANE, J.W. THE BMD AND BMDP SERIES OF STATISTICAL COMPUTER PROGRAMS. COMMUNICATIONS OF THE ACM, OCTOBER 1976, 19(10), 570-576.

GOTT, C.D. MIER-GRP: A COMPUTER PROGRAM FOR THE HIERARCHICAL GROUPING OF REGRESSION EQUATIONS. AFHRL-TR-78-19, AD-A058915. BROOKS AFB, TX: COMPUTATIONAL SCIENCES DIVISION, AIR FORCE HUMAN RESOURCES LABORATORY, JUNE 1978.

GOTT, C.D., & KOPLYAY, J.B. AUTOMATIC INTERACTION DETECTOR -VERSION 4 (AID)-4 REFERENCE MANUAL ADDENDUM 1. AFHRL-TR -77-30, AD-A042968. BROOKS AFB, TX: COMPUTATIONAL SCIENCES DIVISION, AIR FORCE HUMAN RESOURCES LABORATORY, JULY 1977.

KOPLYAY, J.B., GOTT, C.D., & ELTON, J.H. AUTOMATIC INTERACTION DETECTOR-VERSION 4 (AID)-4 REFERENCE MANUAL. AFHRL-TR-73-17, AD-773 803. LACKLAND AFB TX: PERSONNEL RESEARCH DIVISION, AIR FORCE HUMAN RESOURCES LABORATORY, OCTOBER 1973.

MOONAN, W.J. ABCD: A BAYESIAN TECHNIQUE FOR MAKING DISCRIMINATIONS WITH QUALITATIVE VARIABLES. IN PROCEEDINGS FOR 14TH ANNUAL CONFERENCE OF THE MILITARY TESTING ASSOCIATION, 1972, 214-245.

OWEN, D.B. TABLES FOR COMPUTING BIVARIATE NORMAL PROBABILITIES. ANNALS OF MATHEMATICAL STATISTICS, SEPTEMBER-DECEMBER 1956, 27, 1075-1090.

SHAPIRO, S.S., & FRANCIA, R.S. AN APPROXIMATE ANALYSIS OF VARIANCE TEST FOR NORMALITY. JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION, MARCH 1972, 67, 215-216.

SONQUIST, J.A., & MORGAN, J.N. THE DETECTION OF INTERACTION EFFECTS. SURVEY RESEARCH CENTER, INSTITUTE FOR SOCIAL RESEARCH, UNIVERSITY OF MICHIGAN, ANN ARBOR, 1969.

STATJOB SERIES OF REFERENCE MANUALS FOR UNIVAC 1100 SERIES COMPUTERS, ACADEMIC COMPUTING CENTER, UNIVERSITY OF WISCONSIN, MADISON, WISCONSIN.

THE SERIES INCLUDES THE FOLLOWING MANUALS:

INTRODUCTION TO STATJOB, VERSION 10
STATJOB SUMMARY, VERSION 10
COLFREQ1, SINGLE COLUMN FREQUENCY COUNTS, DECEMBER, 1971
CROSTAB2, DATA TABULATION, MARCH, 1975
DISCRIM1, DISCRIMINANT ANALYSIS, MARCH, 1975
DSTAT2, DESCRIPTIVE STATISTICS AND CORRELATION, SEPTEMBER, 1973

FACTOR3, FACTOR ANALYSIS, MAY, 1979
HWAY; FOR PSYCHOLOGISTS, MAY, 1971
ONEWAY2, ONE-WAY ANALYSIS OF VARIANCE, NOVEMBER, 1975
PICT1, SCATTERPLOTS, GRAPHS, AND TABLES, SECOND REVISION, APRIL, 1976
REGAN2, MULTIPLE LINEAR REGRESSION ANALYSIS, APRIL, 1971
ROTATE1, FACTOR ROTATION, NOVEMBER, 1979
STEPREG1, STEPWISE REGRESSION ANALYSIS, AUGUST, 1973
TRANS1, DATA TRANSFER, FIRST REVISION, DECEMBER, 1979
UNISTAT2, UNIVARIATE DESCRIPTIVE STATISTICS AND HISTOGRAMS, APRIL, 1976
BASIC SYSTEM PROCEDURES, MAY, 1970
EXTENDED SYSTEM PROCEDURES, FEBRUARY, 1972
STJBANK PROCEDURES, OCTOBER, 1979
TRANSFORMATION PROCEDURES, OCTOBER, 1975.

STECK, G.P. A TABLE FOR COMPUTING TRIVARIATE NORMAL PROBABILITIES. ANNALS OF MATHEMATICAL STATISTICS, SEPTEMBER-DECEMBER 1958, 29, 780-800.

TUCKER, L.R. MAXIMUM VALIDITY OF A TEST WITH EQUIVALENT ITEMS. PSYCHOMETRIKA, 1946, 21, 1-13.

UNIVAC SYSTEMS REFERENCE LIBRARY. SPERRY UNIVAC 1100 SERIES FUNCTIONAL MATHEMATICAL PROGRAMMING SYSTEM (FMPS) PROGRAMMER REFERENCE. UP-8198. SPERRY RAND CORPORATION, 1975.

UNIVAC SYSTEMS REFERENCE LIBRARY. UNIVAC LARGE SCALE SYSTEMS MATH-PACK PROGRAM ABSTRACTS. UP-4051 REV. 2, SPERRY RAND CORPORATION, 1970.

UNIVAC SYSTEMS REFERENCE LIBRARY. UNIVAC LARGE SCALE SYSTEMS MATH-PACK PROGRAMMERS REFERENCE. UP-7542 REV. 1, SPERRY RAND CORPORATION, 1970.

UNIVAC SYSTEMS REFERENCE LIBRARY. UNIVAC 1106 SYSTEM/1108 MULTI-PROCESSOR SYSTEM STAT-PACK PROGRAM ABSTRACTS. UP-4041 REV. 2, SPERRY RAND CORPORATION, 1969.

UNIVAC SYSTEMS REFERENCE LIBRARY. UNIVAC LARGE SCALE SYSTEMS STAT-PACK PROGRAMMERS REFERENCE. UP-7502 REV. 1, SPERRY RAND CORPORATION, 1970.

VELDMAN, D.J. FORTRAN PROGRAMMING FOR THE BEHAVIORAL SCIENCES. NEW YORK: HOLT, RINEHART AND WINSTON, 1967.

WARD, J.H., JR., BUCHHORN, J., & HALL, E. INTRODUCTION TO PERSUB. PRL-TR-67-3(I), AD-660 578. LACKLAND AFB TX: PERSONNEL RESEARCH LABORATORY, AEROSPACE MEDICAL DIVISION, AUGUST 1967.

WARD, J.H., JR., HALL, E., & BUCHANAN, J. PERSUB REFERENCE MANUAL. PRL-TR-67-21111. AD-660 579. LACKLAND AFB TX: PERSONNEL RESEARCH LABORATORY, AEROSPACE MEDICAL DIVISION, AUGUST 1967.

BIBLIOGRAPHY

ALLEN, J.R. COMPARISON OF STATJOB AND SPSS AND NOTES ON OTHER STATISTICAL SYSTEMS. TECHNICAL PAPERS. USC INCORPORATED, APRIL 1976. 819-832.

HARRIS, R.J. A PRIMER OF MULTIVARIATE STATISTICS. NEW YORK: ACADEMIC PRESS, 1975.

KOONS, R.F., OTAN, T.R., JR., & SELLEMAN, P. INDEX OF AVAILABLE STATISTICAL SOFTWARE. IN PROCEEDINGS OF THE STATISTICAL COMPUTING SECTION, AMERICAN STATISTICAL ASSOCIATION, 1977. 283.